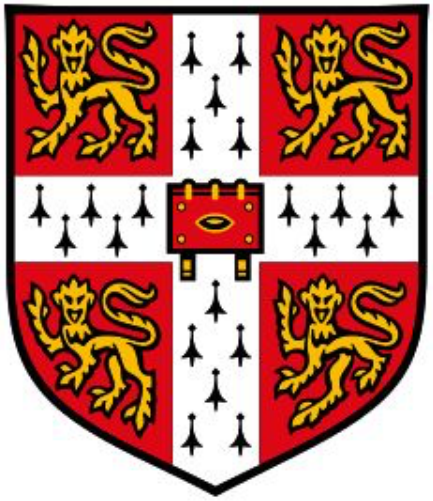
Fantastic ribonucleoprotein complexes and how to study them

Architectural principles of regulatory assemblies in bacteria

Tom Dendooven

A dissertation submitted for the degree of

Doctor of Philosophy



Department of Biochemistry

Churchill college

University of Cambridge

September 2020

Summary

The fate of a bacterial RNA transcript is controlled by three key players: ribonucleases, RNA binding proteins (RBPs) and small regulatory RNAs (sRNAs). Together these form an information-rich regulatory network that exercises post-transcriptional control of gene expression. A key RBP in bacteria is Hfq, which is best studied as a facilitator for sRNA-mRNA pairing. In *Pseudomonas aeruginosa*, however, Hfq is known to modulate translation of more than 100 genes in cooperation with an effector protein, Crc, by sequestering the ribosome binding site of the corresponding mRNA molecules. In this thesis, Cryo-EM studies on target RNAs such as *amiE* and *rbsB,* reveal how Hfq presents multiple A-rich motifs in the mRNA 5′-end to Crc and triggers assembly of higher order complexes. The structures suggest that secondary structure elements and the sequence of the target RNA determine the quaternary structure of the Hfq-Crc translation-repression assembly, introducing a degree of structural diversity that is accommodated by polymorphic interaction surfaces on both Hfq and Crc.

In *Escherichia coli,* the homolog of the same RBP (Hfq) can modulate sRNA lifetime by repurposing the conserved ribonuclease PNPase (polynucleotide phosphorylase) into an RNA chaperone. Cryo-EM structures presented here show how the flexible KH-S1 portal of PNPase generally guides the 3′-end of the sRNA sponge 3′ETS^{leuZ} to the catalytic core for digestion but reroutes the sRNA when bound to Hfq. In particular, an A-rich motif in the 5′-end of 3′ETS^{leuZ} is presented to the KH and S1 domains by the Hfq distal side, while the 3′-end is sequestered on the proximal side. Gel shift- and activity assays suggest that these repurposed 'RNA carrier' complexes protect sRNAs from other ribonucleases such as RNase E and RNase III, and potentially lend their KH-S1 portal as an interaction platform to promote sRNA/RNA target pairing.

A fraction of PNPase in the cell associates with the endoribonuclease RNase E to form the *E. coli '*RNA degradosome', a multi-enzyme ribonuclease assembly responsible for the bulk RNA turnover in the cell. Other partner enzymes include the DEAD-box helicase RhlB and Enolase. The core of the assembly, RNase E, has a long C-terminal scaffold domain that has a conserved natively unstructured character. The subcellular localization of the RNA degradosome, anchored to the cell membrane, provides a spatial component to post-transcriptional gene regulation. In this work, holistic experimental strategies to study this molecular machine are developed. Reproducible reconstitution of a truncated version of the RNA degradosome was achieved with a mix of detergents. Using small angle X-ray solution scattering (SAXS) combined with conformational ensemble computation indicates that the truncated degradosome is a highly flexible system with extended scaffold domains but can adopt more compact conformations when 9S rRNA is bound. Complementary cryo-EM studies of the catalytic core and subsequent 3D variability analyses reveal three modes of molecular motion of the RNase E protomers, adding to the overall flexibility of the RNA degradosome. Lastly, cryo-electron tomography studies of the truncated degradosome tethered to lipid vesicles provide the first visualization of the degradosome in its native environment and hint towards a degree of structural order in its components. Subsequent sub-tomogram averaging approaches to investigate putative membrane-bound RNA surveillance complexes set a foundation for future structural studies.

Acknowledgements

It is somewhat peculiar that one of the first sections of a PhD thesis, the acknowledgements, is usually the last piece to be written. “Last, but not least” as they say. ‘Starting’ this thesis, all that rests is to reflect on all the fantastic people that helped shape this work and acknowledge them accordingly.

First and foremost, I would like to express my immense gratitude to Ben for his support, enthusiasm, and patience. I will certainly miss our brain storming sessions to come up with new and exciting experiments. In particular, I would like to thank Ben for the almost unconditional freedom he gave me to explore new strategies when studying difficult biological systems. Lastly, I truly enjoyed our numerous scientific trips together. It is very stimulating to have a supervisor with such contagious enthusiasm and passion for science. Second, I am grateful to AstraZeneca for funding me through the course of my doctoral studies. Third, my gratitude goes out to the cryo-EM facility, run by Dima, Steve and Lee. I would like to thank Dima for all his help with setting up the many data collections, for the good jokes, and for tolerating my organisational mischief. A special thanks goes out to Steve, my first scientific supervisor, for the excellent advice and all the help early on in the PhD. I would like to thank the Luisi group, Xue, Saif, Kotryna, Miao, Andrzej, Heather, Laura, Tai, Angela, Steven and Yvette, for the fun times in- and out of the lab and for creating a collegiate atmosphere. In particular, I would like the thank Kasia for the amazing collaborations, for teaching me how to work properly in the fragile world of RNA biology, and for the many good laughs we had. I was very fortunate to supervise and work with many talented Master students I: Kou, Oscar, Betka, Giulia, Zak and Marta. Lastly, I was lucky enough to meet some great new friends along the way and would like to thank Roger for his friendship and for inviting me to the John Innes Centre, Sam for providing a welcome distraction with the occasional pub visit and for the word ‘anthology’, and Andrew and Liam for the deceitful yet enjoyable ‘round of rounds’. I am also thankful to Churchill College for the welcome atmosphere they provided and the financial support towards my hobbies.

This thesis would not have been possible without the help of many collaborators in and outside Cambridge. I would like to thank Dr. Jamie Blaza for his helpful advice on cryo-EM, Louis Wilson for the exciting collaboration on EXTL3, Alister Burt for his help with writing tomography scripts, Dr. Alex Shkumatau for his invaluable help with the SAXS ensemble modelling, Prof. Udo Bläsi and Dr. Elisabeth Sonnleitner for the fruitful collaboration on the Hfq-Crc work, Dr. Katherine Stott for the help and advice with Biophysical experiments, Prof. Nicholas De Lay for the fantastic collaboration on the PNPase project and Prof. Jiří Šponer and Dr. Miroslav Krepl for the Hfq-Crc Molecular Dynamics simulations.

Lastly, I am incredibly grateful for the good care and love I received from my mom and dad, my brothers, Steven and Gert, and my parents in law, Marc and Marleen. Thank you, my sweet Ilona, for being there for me and for your unconditional support and love along this journey. I am one lucky guy.

Declaration

I declare that this thesis is the result of my own work, carried out at the Department of Biochemistry, University of Cambridge, between October 2016 and September 2020 and includes nothing which is the outcome of work done in collaboration except as specified explicitly in the text. It is not substantially the same as any work that has already been submitted before for any degree or other qualification except as declared in the preface and specified in the text. It does not exceed the prescribed word limit for the Biology Degree Committee.

Tom Dendooven

Department of Biochemistry

Churchill College

University of Cambridge

September 2020

Contents

List of abbreviations

2xYT	Yeast extract tryptone medium
3DVA	3D variability analysis
A	Adenine
Å	Angstrom = 10^{-10} m
A260	Absorbance at 260 nm wavelength
A280	Absorbance at 280 nm wavelength
AR2	Second arginine-rich region of RNase E, residues 796-814
ATP	Adenosine triphosphate
AUC	Analytical ultracentrifugation
β-DDM	*n*-Dodecyl β-D-maltoside
bp	Base pairs
BSA	Bovine serum albumin
CHAPSO	3-([3-Cholamidopropyl]dimethylammonio)-2-hydroxy-1-propanesulfonate
Cryo-EM	Cryo-electron microscopy
Cryo-ET	Cryo-electron Tomography
ChiP	Chitoporin precursor
CTD	Carboxy-terminal domain
Da	Dalton, the mass of a hydrogen atom
DLS	Dynamic Light Scattering
DNA	Deoxy-ribonucleic acid
dNTP	Deoxy nucleoside triphosphate
DTT	Dithiothreitol
EMSA	Electrophoretic mobility shift assay
ETS	External transcribed Spacer
GO	Graphene oxide
HEPES	N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulphonic acid)
Hfq	Host factor I for phage Qβ, an RNA chaperone protein
IPTG	Isopropyl β-D-thiogalactopyranoside
kDa	kilo Dalton
LB	Luria Bertani broth
M	Molar
sRNA	Small regulatory RNA
MAH	Membrane attachment helix
miRNA	microRNA
mRNA	Messenger RNA

MW	Molecular weight
Nt	Nucleotide(s)
NTD	Amino-terminal domain
OD	Optical density
OmpX	Outer membrane protein X
P	Phosphate
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PNPase	Polynucleotide phosphorylase
RBD	RNA-binding domain of RNase E, residues 604-644
RBS	Ribosome binding site
RhlB	RNA helicase B
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNAP	RNA Polymerase
RNase	Ribonuclease
rNDP	(ribose) nucleoside diphosphate
rNTP	(ribose) nucleoside triphosphate
rRNA	Ribosomal RNA
SAXS	Small-angle X-ray scattering
SEC	Size-exclusion chromatography
SDS	Sodium dodecyl sulphate
siRNA	Short interfering RNA
SPA	Single particle analysis
sRNA	Small regulatory RNA
STAC	Sub-tomogram averaging and classification
TBE	Tris-borate-EDTA buffer
TCEP	Tris(2-carboxyethyl)phosphine
Tris	Tris(hydroxymethyl)aminoethane
tRNA	Transfer RNA
U	Uracil
UTR	Untranslated region
UV	Ultraviolet light
Wt	Wild type

Preface

Table of Contents

1 Riboregulation in Bacteria, a versatile network of molecular relationships

Not too long ago, bacterial gene expression was thought to be regulated entirely at the RNA transcription level. The RNA pool was therefore considered to be relatively minimalistic: transfer RNAs (tRNA), ribosomal RNAs (rRNA) and messenger RNAs (mRNAs) were the core players of a simple translation-centred system. If bacteria transcribed more mRNA, more protein would be produced, and vice versa. Extensive RNA-mediated post-transcriptional control mechanisms such as the Eukaryotic miRNA and RNA interference pathways were not predicted to exist in bacteria (Gorski *et al.*, 2017).

This notion certainly has changed. Over the last two decades extensive bacterial riboregulatory networks have been revealed and were found to rival Eukaryotic post-transcriptional regulation mechanisms in terms of complexity. Firstly, a great deal of non-canonical RNA species was characterized, such as RNA thermometers, riboswitches, CRISPR-Cas RNAs and small-noncoding RNAs (sRNAs) (Kortmann & Narberhaus, 2012; Sherwood & Henkin, 2016; Hille *et al.*, 2016; Wagner & Romby, 2015). Each of these respond to internal (cellular) and external (environmental) triggers to coordinate gene expression, either directly or indirectly, and often help control bacterial phenotypes such as virulence or stress-response pathways (Storz *et al.*, 2011; Gottesman & Storz, 2011). Strikingly, RNA interactome studies in E*scherichia coli* and *Salmonella typhimurium* indicate that no less than half of all mRNAs are subject to sRNA-mediated regulation (Hör & Vogel, 2017).

Secondly, numerous conserved RNA binding proteins (RBPs) have been discovered in bacteria and were shown to be equally important players in post-transcriptional gene-regulation. RBPs like Hfq, ProQ and CsrA often work in conjunction with sRNAs but can also modulate the fate of mRNA targets directly (Updegrove *et al.*, 2016; Olejniczak & Storz, 2017; Wagner & Romby 2015; Pei *et al.*, 2019; Sonnleitner *et al.*, 2018, Dendooven & Luisi, 2017). Hfq in particular is a conserved, well-studied and pleiotropic RNA chaperone in bacteria. For many sRNAs Hfq is required for intracellular stability and often facilitates annealing of the sRNA seed region to RNA targets (De Lay *et al.*, 2013; Santiago-Frangos & Woodson, 2018). These 'Hfq-dependant' sRNAs in turn affect the translation and stability of target RNAs. Translation is regulated by binding the Shine-Dalgarno sequence in the mRNA 5' UTR (untranslated region), thereby controlling access to ribosomes (Papenfort & Vanderpool, 2015). mRNA stability is affected through recruitment of decay machinery by a pairing sRNA or through inhibition of its cleavage activity (Lalaouna *et al.*, 2013; Fröhlich *et al.*, 2013; Hui *et al.*, 2014).

Thirdly, the control of mRNA lifetime plays an essential role in the regulation of gene expression. mRNAs are rapidly degraded in bacteria, often together with a pairing sRNA, allowing for quick responses to changes in cellular status or the environment. Unlike DNA, RNA is therefore inherently labile and synthesised continuously. Although transcription is important to ensure adequate steady-state RNA levels, post-transcriptional control of the RNA pool has been increasingly appreciated as a key regulon as well. In addition, many nascent RNAs are premature as they spool of the RNA polymerase, including non-coding regulatory RNAs, and require processing to be active. Both control of steady-state levels of the RNA pool and RNA processing, but also quality control of RNA transcripts, is carried out by various ribonucleases (RNases) in the cell (Arraiano *et al.*, 2010; Bechhofer & Deutscher, 2019). The most important of these is RNase E, the primary decay-initiating ribonuclease in many Gram-negative bacteria, including *E. coli* (Bandyra & Luisi, 2018). Besides bulk RNA turnover, RNase E is responsible for the maturation and degradation of rRNAs and the maturation of tRNAs (Li *et al.*, 1999b, Sulthana *et al.*, 2016). Notably, RNase E forms the core unit of the canonical RNA degradosome in Gram-negative bacteria (Vanzo *et al.*, 1998). The RNA degradosome is a multienzyme assembly that typically consists of one or more ribonucleases, an RNA helicase and often a metabolic enzyme (Aït-Bara & Carpousis, 2015). It is likely that these degradosome components digest substrate RNAs in a highly cooperative manner (Bandyra *et al.*, 2013). Unfortunately, structural elucidation of the RNA degradosome as a whole is particularly challenging, due to its architecture, and insights on how different enzymes in the assembly cooperate mainly come from studies of smaller sub-assemblies (Bruce *et al.*, 2018; Mildenhall *et al.*, 2016). Consequently, a blueprint of the working mechanism of this integrated molecular machine has remained elusive so far.

Riboregulation in bacteria is thus constituted by a complex, intertwined network of ribonucleases, regulatory sRNAs and RNA chaperones. Ribonucleases can be regarded as the executive engines of the network, guided by sRNAs and facilitated by RNA chaperones. This thesis aims to explore some fascinating nodes in these networks from a structural point of view. As such, the experimental work is written as an anthology of stories

about ribonucleases, sRNAs and RNA chaperones and how they cooperate in often remarkable ways to act on RNA targets. In the next section each story will be introduced briefly.

2 Thesis outline

The thesis is divided in three large chapters, each presenting a different story about post-transcriptional regulation in bacteria. The outset was to optimise and use structural biology approaches to investigate the roles and working mechanisms of key ribonucleoprotein complexes in two Gram-negative bacteria, the model organism *E. coli* and the opportunistic pathogen *Pseudomonas aeruginosa*. A short outline of each chapter is given here.

The first chapter, **Architectural principles of Hfq-Crc mediated gene-regulation**, describes a versatile molecular off-switch for a group of genes in *Pseudomonads*. This regulon allows dangerous pathogens such as *P. aeruginosa* to adequately turn off unnecessary genes in response to nutrient availability in the environment, a process called Carbon Catabolite Repression (CCR) (Rojo, 2010). Notably, the molecular switch operates by suppressing translation of target mRNAs and is constituted by the RNA chaperone Hfq and a partner molecule Crc (Sonnleitner *et al.*, 2018). Together these sequester the ribosome binding site of target mRNAs and prevent translation. The Chapter opens with initial cryo-EM structures of Hfq-Crc complexes on a short recognition fragment of an mRNA target, *amiE*. These structures reveal how Hfq presents the target motif to Crc to form higher order assemblies. The quaternary structure of the Hfq-Crc translation repression assembly is revealed in follow-up structures with the entire 5' region of *amiE*. A recent study found that more than 100 different genes are coregulated by Hfq and Crc in *P. aeruginosa*, which is in agreement with the many phenotypes that are associated with Hfq-Crc mediated gene regulation, including cell motility, biofilm formation and virulence (Kambara *et al.*, 2018; Pusic et al., 2016; Heitzinger, 2016; Huang *et al*., 2012). A selection of genes from this target pool were investigated further in this chapter. Cryo-EM structures of the Hfq-Crc translation repression complex on one of these targets, *rbsB*, reveals that the quaternary structure of the translation repression assembly is likely to be different for every RNA target, yet governed by the same basic interactions. In summary, the research presented in this Chapter captures the architectural principles that enable two small regulatory proteins to form polymorphic translation-repression complexes on many different targets while achieving sequence specificity. Due to their important role in nutrient utilisation and virulence, the Hfq-Crc assemblies may be efficacious targets for antibacterial treatments against *P. aeruginosa*, and the high-resolution structures described here provide an entry point for rationalising design of inhibitory compounds.

The second chapter, **Cooperative PNPase-Hfq carrier complexes protect sRNAs and facilitate handover**, describes how the ancient ribonuclease PNPase (polynucleotide phosphorylase) can be repurposed to fulfil chaperoning roles in *E. coli*. PNPase is generally described as a conserved exoribonuclease that contributes to RNA degradation and quality control in diverse organisms (Cameron *et al.*, 2018). Paradoxically, recent studies highlight that PNPase can adopt a secondary mode of action where it stabilises sRNA species rather than

degrade them, and contributes to sRNA mediated gene regulation (De Lay & Gottesman, 2011; Bandyra *et al.*, 2016). The trigger for PNPase to switch to its protective mode is presentation of the sRNA by Hfq, and the research described in this chapter originates from these peculiar findings. In particular, Cryo-EM structures reveal how the RNA chaperone Hfq can form an RNA carrier complex with PNPase and a sRNA target, which is rerouted away from the PNPase active site. Notably, the structures reveal that the flexibly tethered RNA-binding domains within PNPase form a dynamic platform for its pleiotropic functions. Lastly, sRNA protection by RNA carrier complexes and potential downstream effector roles for these assemblies are explored.

The third chapter, **Un-structural Biology of the bacterial RNA degradosome**, explores different approaches to study the quaternary organisation of the *E. coli* RNA degradation machinery. The RNA degradosome was first discovered by Carpousis *et al.* (1994) and is a multi-enzyme assembly that is a crucial player in RNA turnover and processing, i.e. post-transcriptional gene regulation. On paper, the RNA degradosome is a relatively simple machine, with several copies of only four different enzymes participating in the assembly. Interestingly, a series of peculiar degradosome characteristics, such as long natively unstructured scaffold domains and subcellular localisation provide evolutionary advantages to the cell and add extra layers to bacterial riboregulation (Moffit *et al.*, 2016; Hadjeras *et al.*, 2019; De Pristo *et al.*, 2005). The same conserved degradosome characteristics have rendered structural studies of the assembly as a whole nearly impossible so far. Reproducible reconstitution- and purification protocols for the RNA degradosome are the first step in overcoming these challenges and are the starting point of Chapter III. Next, structural findings gathered over the past 20 years are combined in a SAXS-based integrative structural biology approach to explore the conformational landscape of the degradosome assembly. Finally, Cryo-EM is used to gauge whether association with biologically relevant partner complexes, such as the 30S small ribosomal subunit, or reconstitution on a lipid membrane change the conformational behaviour of the RNA degradosome. In particular, the promising prospects of Cryo-Electron tomography studies and sub-tomogram averaging to ultimately reveal the intrinsic working mechanism of the RNA degradosome are described.

3 Experimental approaches

To study the riboregulatory assemblies described above, cryo-electron microscopy (cryo-EM) was used as the primary experimental technique. Over the past 10 years cryo-EM has become one of the most powerful methods to visualise biomolecules and solve their structures. Today, high-resolution reconstructions of biomolecular complexes ranging from tens of kilo-Daltons to many mega-Daltons are becoming routine, and even whole cells can be imaged. These recent advances can be mainly attributed to significant improvements in both electron detectors and 3D-reconstruction algorithms. A second technique that was used in this thesis is Small Angle X-ray Scattering (SAXS), which is particularly useful for flexible macromolecular samples. A short introduction to both techniques will be given here.

3.1 Cryo-electron microscopy - Single Particle Analysis

Conceptually, cryo-EM is a relatively straightforward technique. In a first step, a sample of the particle of interest is applied on a cryo-EM grid and vitrified in liquid ethane (Figure 1A-C). This vitrification method

results in the preservation of biological specimens at near native condition within a thin film of vitreous ice (Dubochet *et al.*, 1988), and omits the need for 3D-crystals (Figure 1D). The thin ice layer is suspended in small holes in the EM grid, many of which are imaged when collecting a cryo-EM dataset. Ideally, each image contains many single particles in different orientations relative to the electron beam used for imaging (Figure 1E, blue boxes). Lastly, Images of isolated particles are iteratively aligned and reconstructed to obtain a high-resolution 3D average of the biological specimen (Figure 1F).

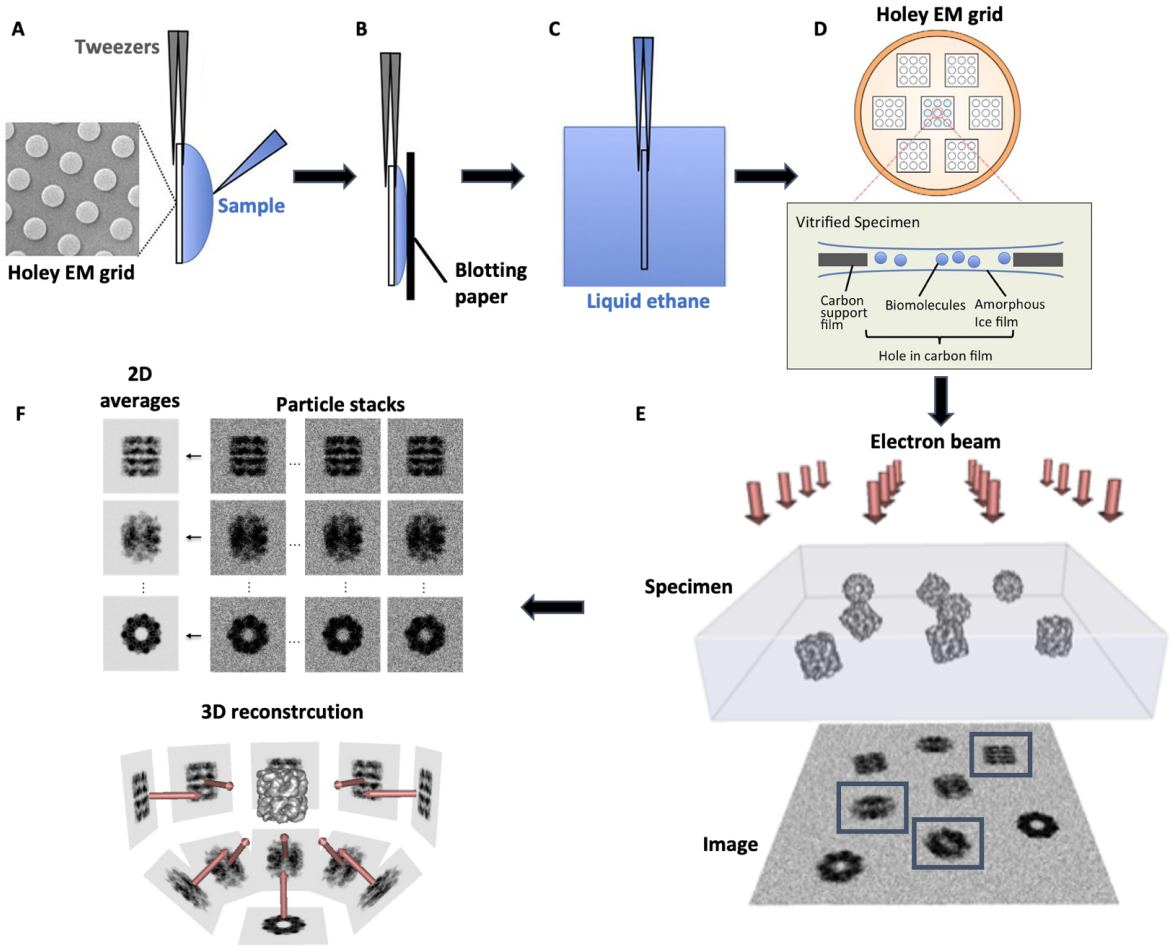


Figure 1: ***Schematic overview of cryo-EM sample preparation, data collection and data processing****. A sample of e.g. a purified protein complex is applied to a holey EM grid* ***(A)****, after which excess sample is removed with blotting paper* ***(B)****. The grid is then vitrified in liquid ethane* ***(C)*** *creating a thin layer of vitreous ice within each grid hole, in which multiple copies of the protein complex are embedded* ***(D)****. The single particles are imaged in an electron microscope* ***(E)*** *and noisy image stacks of the particles are aligned and averaged to increase the signal-to-noise ratio and reconstruct an accurate 3D map of the protein complex* ***(F).*** *Figure adapted from Lyumkis 2019; http://people.csail.mit.edu/gdp/cryoem.html; Murate & Wolf 2018.*

One of the main challenges in analysing cryo-EM images is the low signal-to-noise ratio. Radiation damage of the biological sample by the incident electron beam limits the attainable resolution after 3D reconstruction, and the electron dose must be kept low to minimise this (typically in the range of 40 electrons/Å^2). This in turn leads to high noise levels in the cryo-EM images, with a noise power more than 10 times larger than that of the signal (Nogales & Scheres, 2015). The contrast in the projection image can be described as the sum of

contributions from both amplitude contrast and phase contrast. Biological samples mainly consist of light atoms (H, O, N, and C) that do not absorb electrons from the incident beam. In other words, the total number of electrons in the incident beam and the exit wave immediately after the specimen remains roughly the same, rendering the amplitude contrast in a cryo-EM image negligible. Instead, the specimen scatters the incident electrons at varying angles so that they have different path lengths through the sample, introducing phase contrast in the projection image. Phase contrast is very weak and is usually improved slightly by the combined effects of the spherical aberration of the lens system and collecting images slightly under focus. This combination increases the phase shift between the non-scattered electrons in the incident beam and the scattered electrons in the exit wave. Lastly the poor image contrast can also be increased by using a quarter-wave phase plate, such as the Volta Phase Plate (VPP) which shifts the phase of the scattered electrons in the exit wave relative to that of the transmitted beam by 90° (Danev & Baumeister, 2016). Nevertheless, the low signal-to-noise and poor contrast in cryo-EM images often requires many thousands of projection images to achieve high resolution 3D reconstructions of protein assemblies.

Contrast enhancement by collecting images under focus is standard procedure. However, defocusing complicates downstream operations on the images because features within some resolution bands will have reversed contrast or are have no contrast at all (Figure 2A and B). These projection defects arise from interference of scattered electrons in the exit wave and ultimately constitute an important trade-off: contrast enhancement by defocus versus image distortion. To restore the distorted image, i.e. making sure the 2D-projection corresponds to the electron potential of the real protein assembly, it must be reweighted according to the microscope phase contrast transfer function (CTF) (Figure 2B and C). The CTF describes how the microscope modulates contrast transfer. In practice, the CTF is estimated from the zeros and amplitude signs in the power spectrum (i.e. the Thon rings) of a cryo-EM image, after which the corresponding spatial frequencies are reweighted accordingly (Figure 2C).

The process of defining the relative 3D orientations of the 2D projection images and generating a reconstruction of the biomolecule of interest is called ‘single-particle analysis’ (SPA) (Figure 1E). A series of 2D projections from different orientations can generate a 3D reconstruction by exploiting the ‘projection-slice theorem’, which states that the Fourier transform of a 2D projection of a 3D density is equal to a central slice through the 3D Fourier transform of the density (Figure 2D). Consequently, Fourier transforms of multiple 2D projection images can be combined in Fourier space to produce a 3D Fourier transform of the specimen. In cryo-EM, particles are in random orientations and the relative orientations of the 2D projection images have to be estimated (Figure 2E).

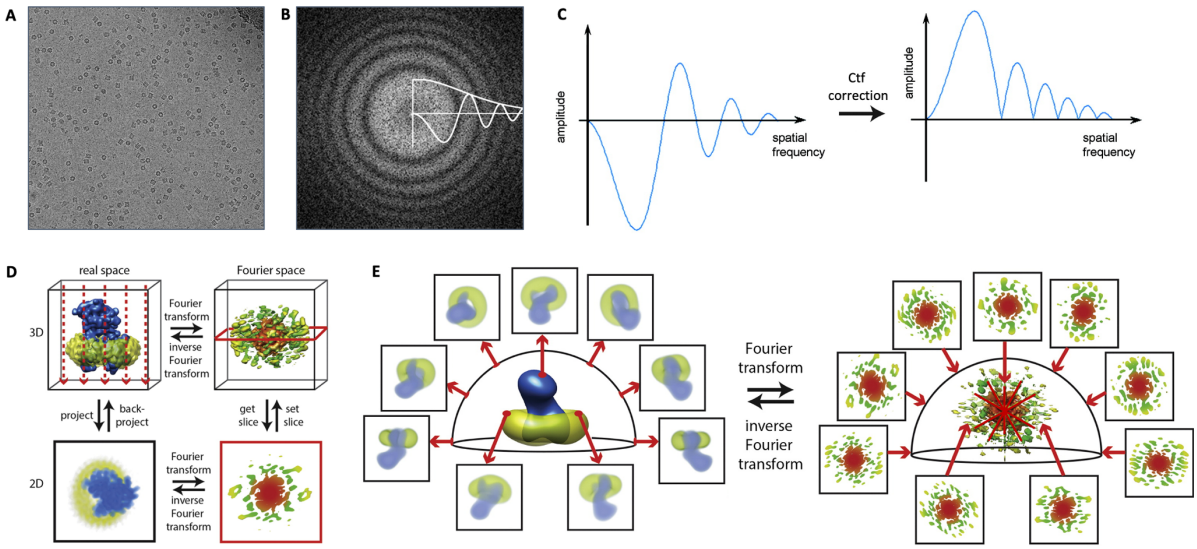


Figure 2: ***CTF estimation and 'projection slice theorem'. (A)*** *Example of cryo-EM micrograph.* ***(B)*** *Power spectrum of the image in (A), with the CTF fitted to the Thon rings.* ***(C)*** *Radial average of the CTF in (B) prior to CTF correction (left) and after CTF correction (phases are flipped, right).* ***(D)*** *Visual representation of the projection-slice theorem. A 3D object (top left) can be projected into a 2D image (bottom left). The 2D Fourier transform of that image (bottom right) corresponds to a central slice through the 3D Fourier transform of the 3D object (top right).* ***(E)*** *If the relative orientations of the 2D projections are known (left), their corresponding Fourier slices can be used to reconstruct the 3D Fourier transform of the object (right). The inverse Fourier transform of this corresponds to the real object. Starting from an initial 3D reference, iterative cycles of (1) calculating the orientations of the 2D projection images and (2) 3D reconstruction to generate a better reference will converge to an accurate model of the 3D object. Images adapted from Nogales & Scheres, 2015; Orlova & Saibil, 2011.*

A popular method to determine the relative orientations of 2D particle images with respect to a 3D reference, i.e. solve the inverse reconstruction problem, is to iteratively compare these to computationally generated projections of the 3D reference structure in different directions, an approach called 'projection-matching' (Penczek *et al.*, 1994). The initial reference structure, which needs to resemble the real structure, is updated in every iteration according to the current orientation angles and converges to the real structure. There are several methods to obtain an initial reference, but the most popular approach is the recent implementation of a stochastic gradient descent-based algorithm (SGD) by Punjani *et al.*, 2017. Here, an initial random reconstruction from a small number of images is incrementally improved by many small and noisy steps. Conceptually, the SGD algorithm performs a wide search in the space of all possible 3D reconstructions to arrive close to the correct structure in the end, albeit at a limited resolution. Major improvements in reconstruction algorithms, such as describing projection orientations as a probability distribution of all possible orientations (maximum likelihood, Scheres *et al.*, 2005; Scheres *et al.*, 2007a) and introducing a regularisation term to reduce the accumulation of noise (Bayesian approach, Scheres, 2012b), helped form the basis for the current cryo-EM revolution (Kühlbrandt, 2014). The final resolution of the calculated reconstruction depends on sample quality, number of images, ice thickness and intrinsic properties of the biomolecule of interest, such as conformational and compositional heterogeneity. The latter two are not merely a challenge for good reconstructions but provide an opportunity to gain insights in molecular mechanisms. Resolving multiple

conformational and compositional states is a common theme in all three chapters of this thesis and recently developed methods to do so will be applied and discussed.

3.2 Cryo-electron tomography and sub-tomogram-averaging

Cryo-electron tomography (cryo-ET) is a variant of the cryo-EM SPA approach described above and has gained significant popularity in the past 10 years, especially for *in situ* structural studies (Beck & Baumeister, 2016). Cryo-ET is particularly useful when the particles of interest are crowded in the vitreous ice layer and/or are part of large superstructures. Cell slices, for example, are packed with many different biomolecules and cellular bodies. The overlapping densities in projection images of such cell slices would make it impossible to identify, annotate or average individual components. The same is true for some *in vitro* systems, for example when protein assemblies are reconstituted on lipid vesicles, as will be the case in Chapter III. In cryo-ET, several projection images are acquired at different angular orientations of the sample, i.e. the microscope stage, resulting in a series of tilted images of the region of interest (Figure 3A). The individual images of a tilt series are then aligned to each other and combined to reconstruct a 3D volume of the specimen, i.e. a tomogram (Figure 3B). To aid fine alignment of the tilt series images, 5-25 nm gold beads are often added to the sample prior to vitrification (Walz *et al.*, 1997). These high-contrast gold fiducials are tracked throughout the tilt series to correct for stage movements in between images and other distortions (Kremer *et al.*, 1996). Alternatively, every image in the tilt series can be segmented in patches that can be tracked throughout the series, omitting the need for gold beads (Amat *et al.* 2010).

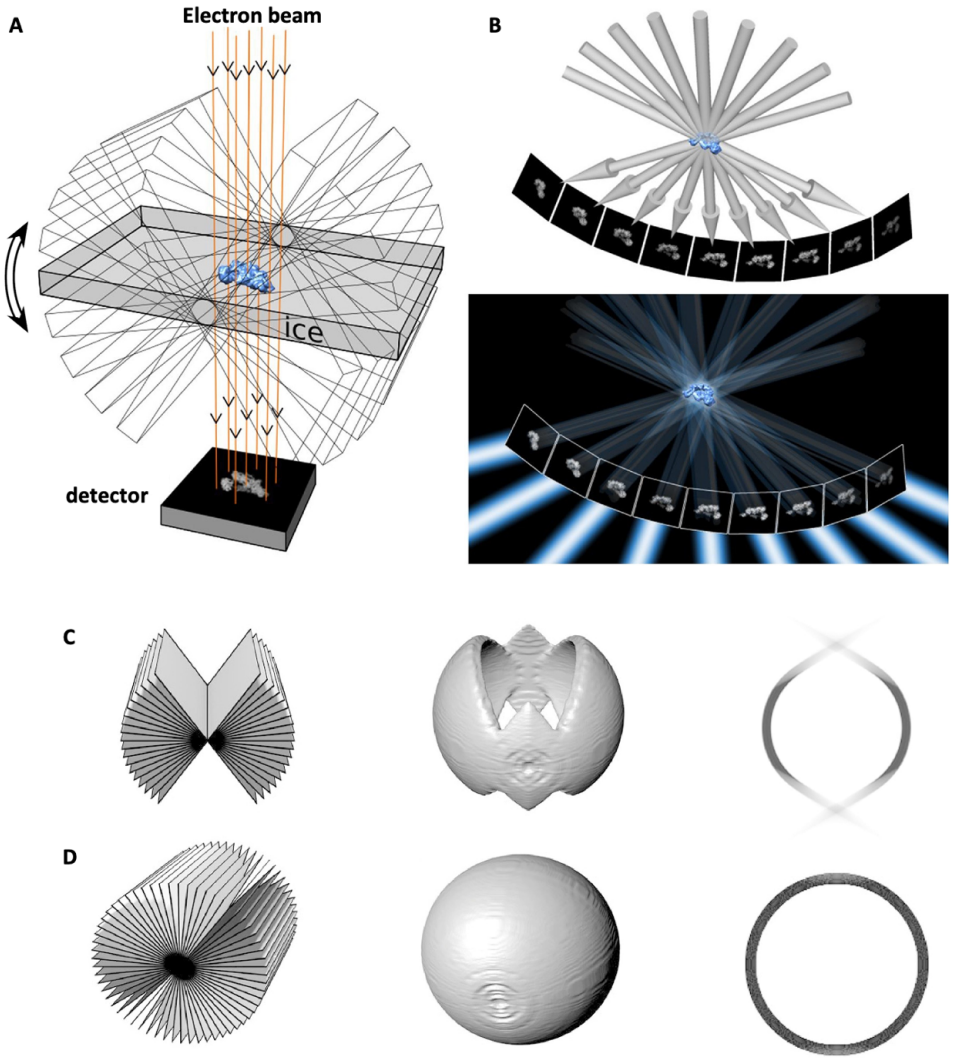


Figure 3: ***Overview of the cryo-electron tomography procedure and missing wedge effects. (A)*** *A series of images is acquired from a sample of interest while the stage is rotated to specified angles. As for SPA, the sample is prepared in a thin layer of vitreous ice but can be very diverse, e.g. whole cells, cellular extracts or membrane-bound proteins.* ***(B)*** *Together, the images form a tilt series containing different projections of the same feature of interest at defined angles.* ***(C)*** *Effects of the missing wedge on a tomographic reconstruction of a spherical object, shown as Fourier slices (left), a reconstructed spherical shell (real space, middle) and a YZ slice through the shell (right).* ***(D)*** *Same as (C) but for full angular sampling along the tilt axis. Images adopted from Koning et al. 2018.*

After tilt series alignment, a 3D reconstruction is generally obtained by weighted back-projection (WBP) of the individual images to account for oversampling of low-resolution components (Radermacher, 1988), much like the reconstruction process described in the previous paragraph (Projection-slice theorem, Figure 2E). In the resulting tomogram, the features that were superimposed in the individual projection images are now separated. Due to the geometry of the EM grids, tilt series can only be collected over a limited range, which is usually set between -60° and +60° (Figure 3C and D). This means that the information content in reconstructed tomograms is incomplete, which manifests in a 'missing wedge' in Fourier space and a reduced resolution along the optical axis at 0° tilt in real space (i.e. the reconstruction is anisotropic in resolution) (Figure 3C and D) (Diebolder *et al.*, 2015). In addition, missing information between sequential tilt images (typically collected

with 2-3° increment) also affects the anisotropic resolution in a tomogram (Koster *et al.*, 1997). On top of these anisotropic resolution effects, the same electron dose limitations as for SPA apply to imaging a tilt series. The limited total dose now needs to be spread over many images of the same area, further decreasing the signal-to-noise ratio, but still has to provide enough signal in each image for accurate CTF-estimation. To balance both requirements, a total dose of approximately 100 electrons/Å^2 per tilt series is typically applied and dedicated tilt schemes are used.

It is clear that raw tomograms are by no means ideal reconstructions. Not only is the resolution limited by the poor signal-to-noise ratio inherent to cryo-EM studies, the limited tilt range results in anisotropic reconstructions (Figure 3C). However, the isotropy as well as the resolution can be recovered for abundant biomolecules in a tomogram by aligning and averaging sub-tomographic volumes containing these. This process, called 'sub-tomogram averaging', is a chimeric approach between classical SPA (2D projections are aligned) and cryo-electron tomography (3D volumes are aligned) and is especially useful for solving macromolecular structures *in situ* (Beck & Baumeister, 2016) or to study membrane-bound assemblies (Sharp *et al.*, 2015), such as the RNA degradosome (Chapter III).

3.3 Small Angle X-ray Scattering

Small angle X-ray scattering (SAXS) is an established biophysical and structural biology technique to investigate the overall oligomerisation state and shape of biological macromolecules in solution. The first SAXS experiments were pioneered by André Guinier in 1938 and have since evolved to the widely applicable method it is today (Guinier, 1938). The macromolecules that can be analysed by SAXS range from small peptides of a few kDa to large complexes of several mDa and even entire viruses, as long as adequate sample concentrations can be achieved. Notably, SAXS is particularly useful for proteins that have domain architecture, i.e. structured domains are connected by flexible linkers.

In a SAXS experiment a macromolecular sample is exposed to high energy photons. The photons are elastically scattering by the electrons in the irradiated sample and are detected to obtain information about the sample's electron density distribution (Figure 4A). To help ensure that the sample is monodisperse, a size-exclusion chromatography (SEC) step is often carried out right before exposure, a method called SEC-SAXS (Mathew *et al.*, 2004). In this setup, the eluent of the SEC column goes straight to the sample chamber and scattering profiles are collected directly after separation of potential aggregates. In solution, the particles of interest have random orientations, which results in isotropic scattering and allows for radial averaging of a 2D detector image (Figure 4A and B). The result is a linear 1D scattering curve where scattering intensities are plotted against the scattering vector (q, ~spatial frequency), often referred to as the intensity profile (Figure 4B). Where cryo-EM datasets are collected in real space, the SAXS intensity profile is recorded in reciprocal space. This means that intensities at low q values in the scattering profile correspond to larger features of the particle of interest, such as overall particle shape and oligomerisation state. Scattering at high q values, on the other hand, presents information about the finer details of the particle structure.

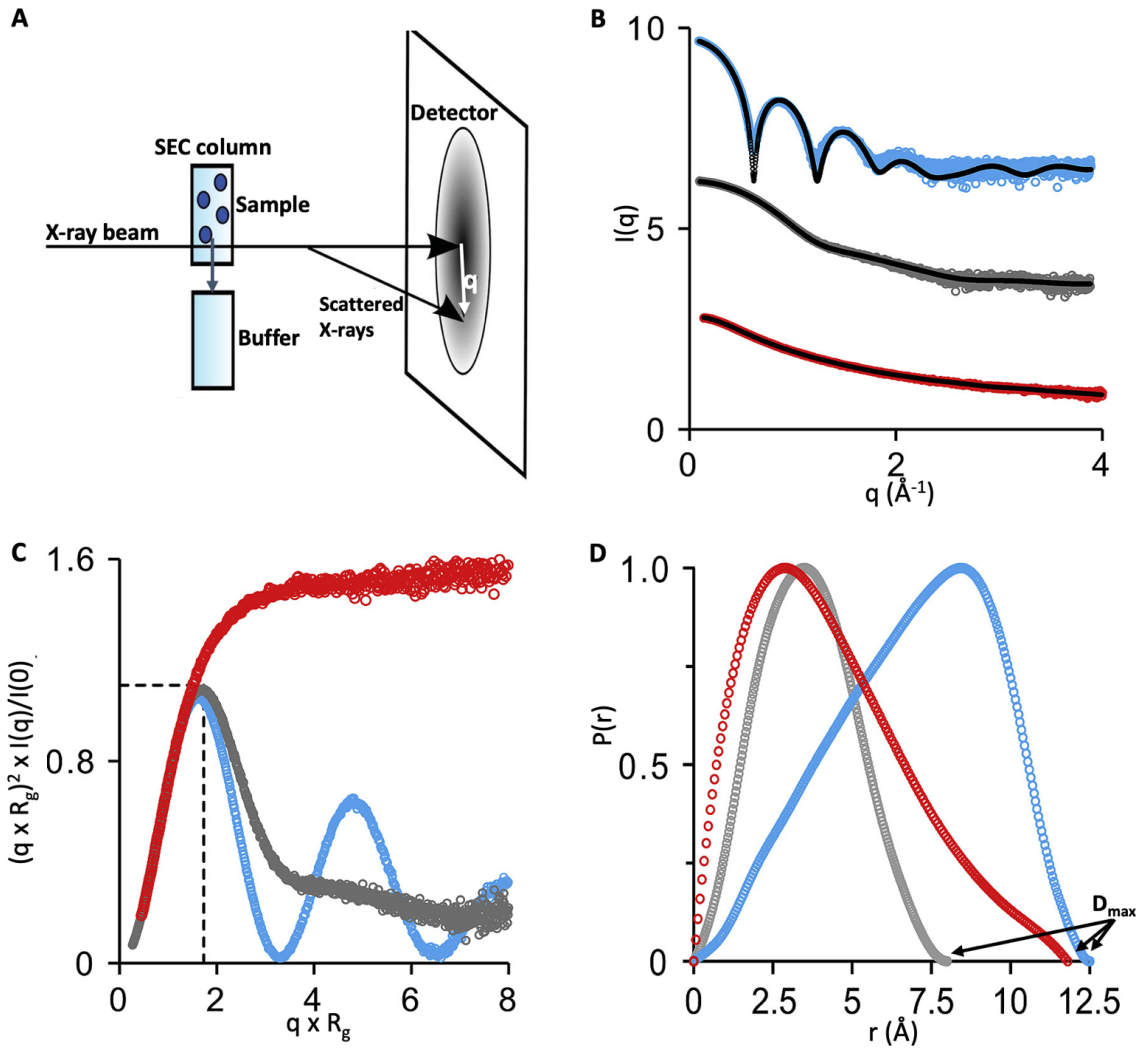


Figure 4: ***Overview of a SAXS experiment and different data presentations. (A)*** *A sample is run over a size exclusion column and irradiated with a collimated X-ray beam in the eluent. Elastically scattered electrons are detected as an isotropic 2D scattering profile on a detector.* ***(B)*** *A 1D intensity curve is obtained by radially averaging the 2D scattering profile and subtracting the intensity curve of the running buffer. Intensity curves are depicted for proteins corresponding to a hollow sphere (blue), a compact full sphere (grey) and a disordered rod (red).* ***(C)*** *The dimensionless Kratky plot reveals the conformational character of the particles. Globular proteins show a clear descent after a (first) maximum (blue and grey curves) whereas disordered proteins show a rising curve that plateaus (red curve).* ***(D)*** *Pairwise distribution functions for the three proteins, showing all the pairwise distances between scattering elements in the particle. The largest dimension of each protein (D_{max}) is annotated.*

From the intensity profile, some general structural parameters of the sample can be calculated. For example, the radius of gyration (R_g), a measure for the particle size, can be derived from the lowest q-regions of the intensity profile. Information about the conformational behaviour of the sample is encoded in the middle q region of the SAXS curve and is visualized by the normalized Kratky plot (Figure 4C) (Kratky & Porod, 1949). Compact and structured objects have a bell-shaped curve in the Kratky plot (Figure 4C, grey curve), whereas unfolded or intrinsically disordered proteins (IDPs) don't have a maximum, but rise to a plateau (Figure 4C, red curve). In between are e.g. particles that are folded but have flexible linkers connecting domains, where the

maximum is shifted to higher q-regions. Lastly, the indirect Fourier transform of the intensity curve generates the electron pairwise distribution function (P(r)) (Figure 4D). The P(r) function is a histogram containing all the pairwise distances between scattering units in the particle and provides rudimentary information about its shape. That is, differently shaped particles, such as solid spheres, hollow spheres or rods will display different features in their P(r) functions (Figure 4D, grey, blue and red curves, respectively). In addition, the maximum intra-particle distance, D_{max}, can be derived from the distribution function.

Perhaps the biggest advantage of SAXS is its ability to easily include *a priori* information about a protein assembly from other structural methods, such as high-resolution X-ray crystallography structures, cryo-EM models, or NMR data. Many biological systems are inherently flexible, meaning that they exist as a large pool of coexisting conformers in solution. A way to deal with such systems in SAXS is through ensemble modelling, a method that was first described by Bernado *et al.* (2007). Ensemble modelling aims to incorporate a priori information about a flexible protein assembly, e.g. domain structures, to find a set of candidate models that, as a mixture, fit the experimental intensity curve. This approach is particularly suited for protein assemblies such as the RNA degradosome and will be described extensively in Chapter III.

4 References

Aït-Bara, S., & Carpousis, A. J. (2015). RNA degradosomes in bacteria and chloroplasts: classification, distribution and evolution of RNase E homologs. *Molecular microbiology*, *97*(6), 1021–1135. https://doi.org/10.1111/mmi.13095

Amat, F., Castaño-Diez, D., Lawrence, A., Moussavi, F., Winkler, H. & Horowitz, M. (2010) Chapter Thirteen - Alignment of Cryo-Electron Tomography Datasets. In: Jensen GJBT-M in E, ed. *Cryo-EM, Part B: 3-D Reconstruction*. Vol 482. Academic Press, 343-367. doi:https://doi.org/10.1016/S0076-6879(10)82014-2

Arraiano, C. M., Andrade, J. M., Domingues, S., Guinote, I. B., Malecki, M., Matos, R. G., Moreira, R. N., Pobre, V., Reis, F. P., Saramago, M., Silva, I. J., & Viegas, S. C. (2010). The critical role of RNA processing and degradation in the control of gene expression. *FEMS microbiology reviews*, *34*(5), 883–923. https://doi.org/10.1111/j.1574-6976.2010.00242.x

Bandyra, K. J., & Luisi, B. F. (2018). RNase E and the High-Fidelity Orchestration of RNA Metabolism. *Microbiology spectrum*, *6*(2), 10.1128/microbiolspec.RWR-0008-2017. https://doi.org/10.1128/microbiolspec.RWR-0008-2017

Bandyra, K. J., Bouvier, M., Carpousis, A. J., & Luisi, B. F. (2013). The social fabric of the RNA degradosome. *Biochimica et biophysica acta*, *1829*(6-7), 514–522. https://doi.org/10.1016/j.bbagrm.2013.02.011

Bandyra, K.J., Sinha, D., Syrjanen, J., Luisi, B.F. and De Lay, N.R. (2016) The ribonuclease polynucleotide phosphorylase can interact with small noncoding RNAs in both stabilizing and degradative modes. *RNA* 22, 360-372. PMC4748814

Bechhofer, D. H., & Deutscher, M. P. (2019). Bacterial ribonucleases and their roles in RNA metabolism. *Critical reviews in biochemistry and molecular biology*, *54*(3), 242–300. https://doi.org/10.1080/10409238.2019.1651816

Beck, M. & Baumeister, W. (2016) Cryo-Electron Tomography: Can it Reveal the Molecular Sociology of Cells in Atomic Detail? *Trends Cell Biol.*, 26(11), 825-837. doi:https://doi.org/10.1016/j.tcb.2016.08.006

Bernado, P., Mylonas, E., Petoukhov, M.V., Blackledge, M., Svergun, D.I. (2007) Structural Characterization of Flexible Proteins Using Small-Angle X-ray Scattering. *J. Am. Chem. Soc.* **129(17)**, 5656-5664

Bruce, H. A., Du, D., Matak-Vinkovic, D., Bandyra, K. J., Broadhurst, R. W., Martin, E., Sobott, F., Shkumatov, A. V., & Luisi, B. F. (2018). Analysis of the natively unstructured RNA/protein-recognition core in the Escherichia

coli RNA degradosome and its interactions with regulatory RNA/Hfq complexes. *Nucleic acids research*, *46*(1), 387–402. https://doi.org/10.1093/nar/gkx1083

Cameron, T.A., Matz, L.M. and De Lay, N.R. (2018) Polynucleotide phosphorylase: not merely an RNase but a pivotal post-transcriptional regulator. PLos Genetics 14:e1007654.

Carpousis, A. J., Van Houwe, G., Ehretsmann, C., & Krisch, H. M. (1994). Copurification of E. coli RNAase E and PNPase: evidence for a specific association between two enzymes important in RNA processing and degradation. *Cell*, *76*(5), 889–900. https://doi.org/10.1016/0092-8674(94)90363-8

Danev, R., & Baumeister, W. (2016). Cryo-EM single particle analysis with the Volta phase plate. *eLife*, *5*, e13046. https://doi.org/10.7554/eLife.13046

De Lay, N., Schu, D. J., & Gottesman, S. (2013). Bacterial small RNA-based negative regulation: Hfq and its accomplices. *The Journal of biological chemistry*, *288*(12), 7996–8003. https://doi.org/10.1074/jbc.R112.441386

De Lay N, Gottesman S. (2011) Role of polynucleotide phosphorylase in sRNA function in *Escherichia coli*. *RNA*, 17(6), 1172–1189.

Dendooven, T., & Luisi, B. F. (2017). RNA search engines empower the bacterial intranet. *Biochemical Society transactions*, *45*(4), 987–997. https://doi.org/10.1042/BST20160373

DePristo, M. A., Weinreich, D. M., & Hartl, D. L. (2005). Missense meanderings in sequence space: a biophysical view of protein evolution. *Nature reviews. Genetics*, *6*(9), 678–687. https://doi.org/10.1038/nrg1672

Diebolder, C.A., Faas, F.G.A., Koster, A.J., Koning, R.I. Conical Fourier shell correlation applied to electron tomograms. *J Struct Biol*., 190(2), 215-223. doi:https://doi.org/10.1016/j.jsb.2015.03.010

Dubochet, J., Adrian, M., Chang, J. J., Homo, J. C., Lepault, J., McDowall, A. W., & Schultz, P. (1988). Cryo-electron microscopy of vitrified specimens. *Quarterly reviews of biophysics*, *21*(2), 129–228. https://doi.org/10.1017/s0033583500004297

Fröhlich, K. S., Papenfort, K., Fekete, A., & Vogel, J. (2013). A small RNA activates CFA synthase by isoform-specific mRNA stabilization. *The EMBO journal*, *32*(22), 2963–2979. https://doi.org/10.1038/emboj.2013.222

Gorski, S. A., Vogel, J., & Doudna, J. A. (2017). RNA-based recognition and targeting: sowing the seeds of specificity. *Nature reviews. Molecular cell biology*, *18*(4), 215–228. https://doi.org/10.1038/nrm.2016.174

Gottesman, S., & Storz, G. (2011). Bacterial small RNA regulators: versatile roles and rapidly evolving variations. *Cold Spring Harbor perspectives in biology*, *3*(12), a003798. https://doi.org/10.1101/cshperspect.a003798

GUINIER, A. (1938) Structure of Age-Hardened Aluminium-Copper Alloys. *Nature,* **142,** 569–570. https://doi.org/10.1038/142569b0

Hadjeras, L., Poljak, L., Bouvier, M., Morin-Ogier, Q., Canal, I., Cocaign-Bousquet, M., Girbal, L., & Carpousis, A. J. (2019). Detachment of the RNA degradosome from the inner membrane of Escherichia coli results in a global slowdown of mRNA degradation, proteolysis of RNase E and increased turnover of ribosome-free transcripts. *Molecular microbiology*, *111*(6), 1715–1731. https://doi.org/10.1111/mmi.14248

Heitzinger, D.A. (2016) Impact of Hfq on antibiotic susceptibility of *Pseudomonas aeruginosa*. Master Thesis, University of Vienna.

Hille, F., & Charpentier, E. (2016). CRISPR-Cas: biology, mechanisms and relevance. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *371*(1707), 20150496. https://doi.org/10.1098/rstb.2015.0496

Hör, J., & Vogel, J. (2017). Global snapshots of bacterial RNA networks. *The EMBO journal*, *36*(3), 245–247. https://doi.org/10.15252/embj.201696072

Huang, J., Sonnleitner, E., Ren, B., Xu, Y., and Haas, D. (2012). Catabolite repression control of pyocyanin biosynthesis at an intersection of primary and secondary metabolism in *Pseudomonas aeruginosa*. Appl. Environ. Microbiol. *78*, 5016-5020.

Hui, M. P., Foley, P. L., & Belasco, J. G. (2014). Messenger RNA degradation in bacterial cells. *Annual review of genetics*, *48*, 537–559. https://doi.org/10.1146/annurev-genet-120213-092340

Kambara, T.K., Ramsey, K.M., and Dove, S.L. (2018). Pervasive targeting of nascent transcripts by Hfq. Cell Reports *23*, 1543-1552.

Koning, R.I., Koster, A.J., & Sharp, T.H. (2018) Advances in cryo-electron tomography for biology and medicine. *Ann Anat - Anat Anzeiger*., 217, 82-96. doi:https://doi.org/10.1016/j.aanat.2018.02.004

Kortmann, J., & Narberhaus, F. (2012). Bacterial RNA thermometers: molecular zippers and switches. *Nature reviews. Microbiology*, *10*(4), 255–265. https://doi.org/10.1038/nrmicro2730

Koster, A.J., Grimm, R., Typke, D., et al. (1997) Perspectives of Molecular and Cellular Electron Tomography. *J Struct Biol*., 120(3), 276-308. doi:https://doi.org/10.1006/jsbi.1997.3933

Kratky, O. and Porod, G. (1949), Röntgenuntersuchung gelöster Fadenmoleküle. Recl. Trav. Chim. Pays-Bas, 68: 1106-1122. doi:10.1002/recl.19490681203

Kremer, J.R., Mastronarde, D.N. and McIntosh, J.N. (1996) Computer visualization of three-dimensional image data using IMOD *J. Struct. Biol.* **116**:71-76

Kühlbrandt, W. (2014) The Resolution Revolution. *Science (80-)*, 343(6178), 1443 LP - 1444. doi:10.1126/science.1251652

Lalaouna, D., Simoneau-Roy, M., Lafontaine, D., & Massé, E. (2013). Regulatory RNAs and target mRNA decay in prokaryotes. *Biochimica et biophysica acta*, *1829*(6-7), 742–747. https://doi.org/10.1016/j.bbagrm.2013.02.013

Li, Z., Pandit, S., & Deutscher, M. P. (1999). RNase G (CafA protein) and RNase E are both required for the 5' maturation of 16S ribosomal RNA. *The EMBO journal*, *18*(10), 2878–2885. https://doi.org/10.1093/emboj/18.10.2878

Lyumkis, D. (2019) Challenges and opportunities in cryo-EM single-particle analysis. *J Biol Chem*., 294(13), 5181–5197.

Mathew, E., Mirza, A., & Menhart, N. (2004). Liquid-chromatography-coupled SAXS for accurate sizing of aggregating proteins. *Journal of synchrotron radiation*, *11*(Pt 4), 314–318. https://doi.org/10.1107/S0909049504014086

Mildenhall, K.B., Wiese, N., Chung, D., Maples, V.F., Mohanty, B.K. and Kushner, S.R. (2016), RNase E-based degradosome modulates polyadenylation of mRNAs after Rho-independent transcription terminators in *Escherichia coli*. Molecular Microbiology, 101: 645-655. doi:10.1111/mmi.13413

Moffitt, J. R., Pandey, S., Boettiger, A. N., Wang, S., & Zhuang, X. (2016). Spatial organization shapes the turnover of a bacterial transcriptome. *eLife*, *5*, e13065. https://doi.org/10.7554/eLife.13065

Murata, K. & Wolf, M. (2018) Cryo-electron microscopy for structural analysis of dynamic biological macromolecules. *Biochim Biophys Acta - Gen Subj*., 1862(2),324-334. doi:https://doi.org/10.1016/j.bbagen.2017.07.020

Nogales, E. & Scheres, S.H.W. (2015) Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity. *Mol Cell*., 58(4,677-689. doi:https://doi.org/10.1016/j.molcel.2015.02.019

Olejniczak, M., & Storz, G. (2017). ProQ/FinO-domain proteins: another ubiquitous family of RNA matchmakers? *Molecular microbiology*, *104*(6), 905–915. https://doi.org/10.1111/mmi.13679

Orlova, E.V. & Saibil H.R. (2011) Structural Analysis of Macromolecular Assemblies by Electron Microscopy. *Chem Rev*., 111(12), 7710-7748. doi:10.1021/cr100353t

Papenfort, K., & Vanderpool, C. K. (2015). Target activation by regulatory RNAs in bacteria. *FEMS microbiology reviews*, *39*(3), 362–378. https://doi.org/10.1093/femsre/fuv016

Pei, X.Y.*, Dendooven, T.*, Sonnleitner, E., Chen, S., Bläsi U, Luisi, B.F. (2019) Architectural principles for Hfq/Crc-mediated regulation of gene expression. eLife; 8:e43158.

Penczek, P.A., Grassucci, R.A. & Frank, J. (1994) The ribosome at improved resolution: New techniques for merging and orientation refinement in 3D cryo-electron microscopy of biological particles. *Ultramicroscopy*, 53(3), 251-270. doi:https://doi.org/10.1016/0304-3991(94)90038-8

Punjani, A., Rubinstein, J. L., Fleet, D. J., & Brubaker, M. A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature methods*, **14**, 290–296. https://doi.org/10.1038/nmeth.4169

Pusic, P., Tata, M., Wolfinger, M.T., Sonnleitner, E., Häussler, S. and Bläsi, U. (2016). Cross-regulation by CrcZ RNA controls anoxic biofilm formation in *Pseudomonas aeruginosa*. Sci. Rep. *6*, 39621.

Radermacher, M. (1988), Three-Dimensional reconstruction of single particles from random and nonrandom tilt series. J. Elec. Microsc. Tech., 9: 359-394. doi:10.1002/jemt.1060090405

Rojo, F. (2010). Carbon catabolite repression in *Pseudomonas*: optimizing metabolic versatility and interactions with the environment. FEMS Microbiol. Rev*. 34*, 658–684.

Santiago-Frangos, A., & Woodson, S. A. (2018). Hfq chaperone brings speed dating to bacterial sRNA. *Wiley interdisciplinary reviews. RNA*, *9*(4), e1475. https://doi.org/10.1002/wrna.1475

Scheres, S.H.W., Valle, M., Nuñez, R., *et al.* (2005) Maximum-likelihood Multi-reference Refinement for Electron Microscopy Images. *J Mol Biol*., 348(1), 139-149. doi:https://doi.org/10.1016/j.jmb.2005.02.031

Scheres, S., Gao, H., Valle, M. *et al.* (2007) Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. *Nat Methods,* 4, 27–29. https://doi.org/10.1038/nmeth992

Scheres, S.H.W. (2012) RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol*., 180(3), 519-530. doi:https://doi.org/10.1016/j.jsb.2012.09.006

Sherwood, A. V., & Henkin, T. M. (2016). Riboswitch-Mediated Gene Regulation: Novel RNA Architectures Dictate Gene Expression Responses. *Annual review of microbiology*, *70*, 361–374. https://doi.org/10.1146/annurev-micro-091014-104306

Sonnleitner, E., Wulf, A., Campagne, S., Pei, X., Wolfinger, M., Forlani, G., Prindl, K., Abdou, L., Resch, A., Allain, F., Luisi, B., Urlaub, H. and Bläsi, U. (2018). Interplay between the catabolite repression control protein Crc, Hfq and RNA in Hfq-dependent translational regulation in Pseudomonas aeruginosa. Nucleic Acids Res. *46*, 1470-1485.

Storz, G., Vogel, J., & Wassarman, K. M. (2011). Regulation by small RNAs in bacteria: expanding frontiers. *Molecular cell*, *43*(6), 880–891. https://doi.org/10.1016/j.molcel.2011.08.022

Sulthana, S., Basturea, G. N., & Deutscher, M. P. (2016). Elucidation of pathways of ribosomal RNA degradation: an essential role for RNase E. *RNA (New York, N.Y.)*, *22*(8), 1163–1171. https://doi.org/10.1261/rna.056275.116

Tria, G., Mertens, H. D. T., Kachala, M. & Svergun, D. I. (2015) Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ* **2**, 207-217

Updegrove, T. B., Zhang, A., & Storz, G. (2016). Hfq: the flexible RNA matchmaker. *Current opinion in microbiology*, *30*, 133–138. https://doi.org/10.1016/j.mib.2016.02.003

Vanzo, N. F., Li, Y. S., Py, B., Blum, E., Higgins, C. F., Raynal, L. C., Krisch, H. M., & Carpousis, A. J. (1998). Ribonuclease E organizes the protein interactions in the Escherichia coli RNA degradosome. *Genes & development*, *12*(17), 2770–2781. https://doi.org/10.1101/gad.12.17.2770

Wagner, E., & Romby, P. (2015). Small RNAs in bacteria and archaea: who they are, what they do, and how they do it. *Advances in genetics*, *90*, 133–208. https://doi.org/10.1016/bs.adgen.2015.05.001

Walz, J., Typke, D., Nitsch, M., Koster, A.J., Hegerl, R., Baumeister, W. Electron Tomography of Single Ice-Embedded Macromolecules: Three-Dimensional Alignment and Classification. *J Struct Biol*.,120(3),387-395. doi:https://doi.org/10.1006/jsbi.1997.3934

Chapter I: Architectural principles of Hfq-Crc mediated gene-regulation

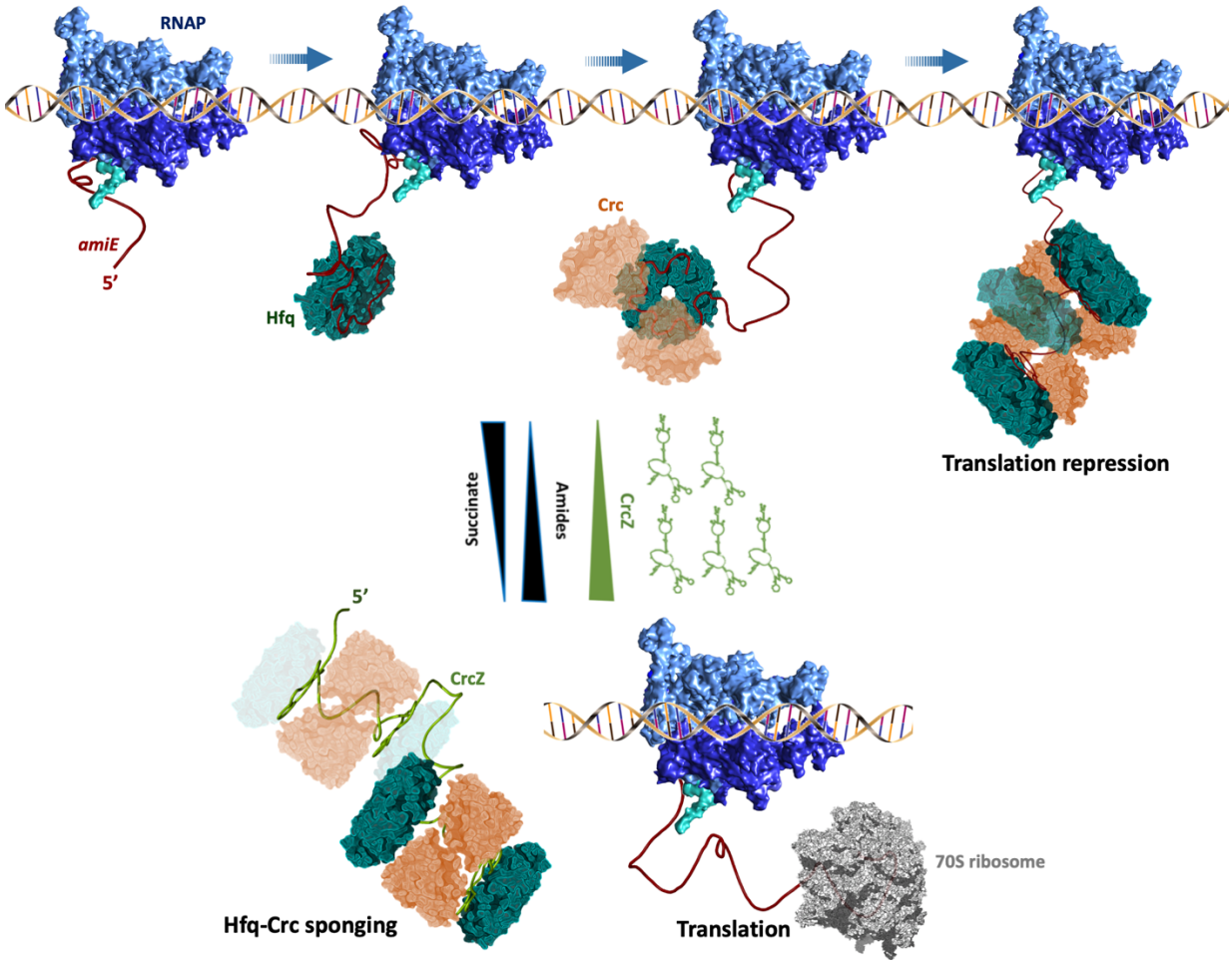


Table of Contents

1 Abstract

In many bacterial species, the global regulator Hfq acts as a key node in post-transcriptional networks that control expression of numerous genes. In the opportunistic pathogen *Pseudomonas aeruginosa*, Hfq controls translation of many target transcripts by forming a regulatory complex with the catabolite repression protein Crc. This translation-repression complex acts as part of an intricate mechanism of preferred nutrient utilisation. In this chapter, high-resolution cryo-EM structures are described of the assembly of Hfq and Crc bound to the translation initiation site of a target mRNA as well as the entire 5' region of the transcript. The core of the assemblies is formed through cooperative interactions between the RNA target, Hfq and Crc. Complex formation on other mRNA targets is further explored by electrophoretic mobility shift assays, and a cryo-EM structure was solved for one of these. In addition, molecular dynamics (MD) simulations of the Hfq-Crc assemblies suggest a mechanism for rapid RNA scanning by Hfq. The findings in this chapter support a versatile assembly model based on interaction degeneracy, explaining how Hfq presents many mRNAs to the same partner protein to form diverse riboregulatory complexes that control translation. In particular, the results described here provide an RNA code that guides global regulators to interact cooperatively and regulate many RNA targets co-transcriptionally.

2 Introduction

Hfq is a conserved bacterial RNA chaperone that is best studied for its contribution to the control of mRNA translation through different modes of action. For one, Hfq can act indirectly by facilitating base-pairing interactions of cognate mRNA targets with small regulatory RNAs (Vogel & Luisi, 2011; Wagner & Romby, 2015). In a second mode, the RNA chaperone can bind A-rich sequences at or in the vicinity of translation initiation sites of mRNA targets to directly repress translation independently of small RNAs (Vecerek *et al.*, 2005; Salvail *et al.*, 2013; Sonnleitner & Bläsi, 2014). As a member of the Lsm/Sm protein family, Hfq shares an ancient structural core that oligomerizes to form a hexameric ring-like structure, exposing several RNA-binding surfaces (Figure 1). RNA recognition is mediated by distinct interactions with distal, proximal and rim faces of the hexameric ring, as shown by crystallographic and biophysical studies (Schumacher *et al.*, 2002; Link *et al.*, 2009; Sauer *et al.*, 2012; Horstmann *et al.*, 2012; Panja *et al.*, 2013). In addition, Hfq has a long, natively unstructured C-terminal tail which participates in auto-regulating RNA-binding activities (Santiago-Frangos *et al.*, 2016; Santiago-Frangos *et al.*, 2017; Santiago-Frangos *et al.*, 2019).

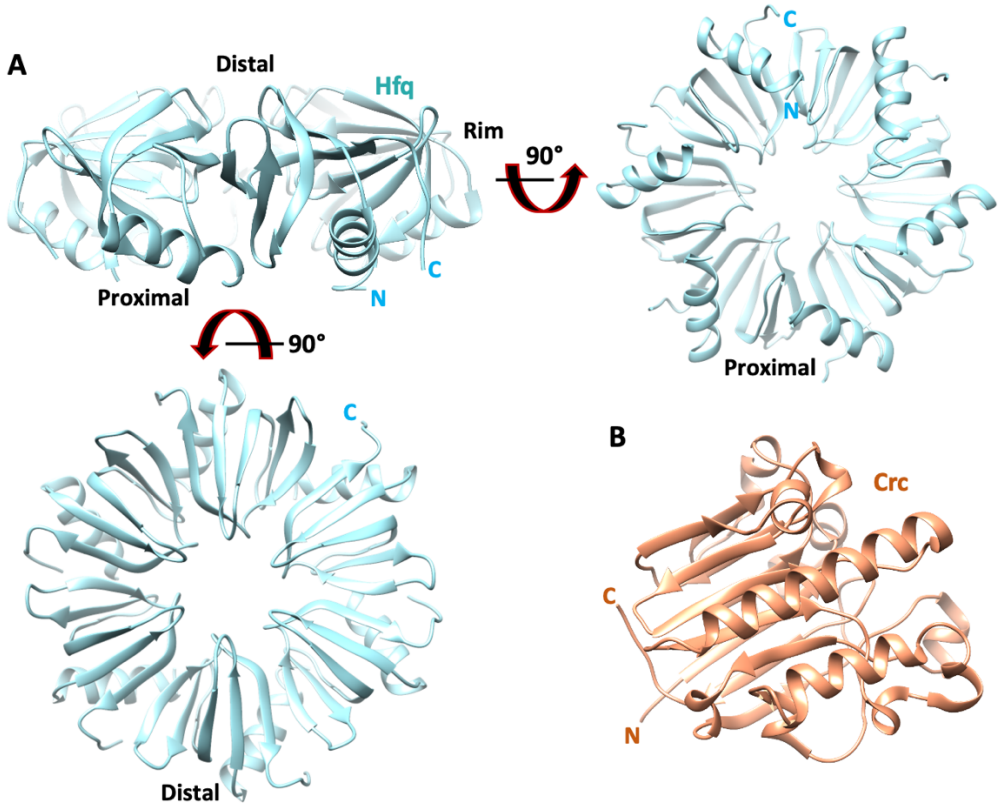


Figure 1: ***Crystal structures of Pseudomonas aeruginosa PAO1 Hfq (cyan) and Crc (orange)****. **(A)** Hfq is a homo-hexameric protein and has a proximal side, which is known to bind 3' poly-uridine tails, a distal side, which binds A-rich sequences and a rim side (pdb-ID 1u1t) **(B)** Crc is a compact, globular monomeric protein with no obvious RNA binding patches exposed on its surface (pdb-ID 4jg3). The N-terminus (N) and C-terminus (C) for Hfq and Crc protomers are annotated.*

Hfq is one of the most pleiotropic RNA chaperones in bacterial riboregulation. In the opportunistic, Gram-negative pathogen *Pseudomonas aeruginosa* Hfq helps coordinate stress responses (Lu *et al.*, 2016), metabolism (Sonnleitner & Bläsi, 2014), quorum sensing (Sonnleitner *et al.*, 2006; Yang *et al.*, 2015) and virulence (Sonnleitner *et al.*, 2003; Fernàndez *et al.*, 2016; Pusic *et al.*, 2016). Many of these roles are likely to be mediated through effector assemblies with Hfq, and several putative protein interactors of *P. aeruginosa* Hfq have been identified with functions in transcription, translation and mRNA decay (Van den Bossche *et al.*, 2014). For Hfq from *Escherichia coli*, for example, the functionally important partners include RNA polymerase, ribosomal protein S1 (Sukhodolets & Garges, 2003), the endoribonuclease RNase E (Ikeda *et al.*, 2011), polyA-polymerase and the exoribonuclease polynucleotide phosphorylase (Mohanty *et al.*, 2004; Bandyra *et al.*, 2016). Most likely, formation of such complexes is at least partly driven by RNA binding, and coordinates the co-localisation of transcription-, translation- and RNA decay machineries in the bacterial cell (Worrall *et al.*, 2008; Resch *et al.*, 2010; Vecerek *et al.*, 2010).

An interesting protein that cooperates with Hfq in *P. aeruginosa* is the catabolite repression control protein, Crc (Figure 1) (Van den Bossche *et al.*, 2014; Moreno *et al.*, 2015; Sonnleitner *et al.*, 2018). Crc is a specialized protein that is found in a limited subset of bacteria including *Pseudomonads* (but excluding *Enterobacteriaceae* such as *E. coli*). The Crc protein is a member of the EEP family (Exonuclease/Endonuclease/Phosphatase), but has lost

enzymatic activity, as indicated by the absence of key catalytic residues in the active site (Milojevic *et al.*, 2013). In *P. aeruginosa*, Crc helps to direct metabolic pathways during carbon catabolite repression (CCR) to use preferential carbon sources, such as succinate, until exhausted, after which alternative carbon sources are used (Rojo, 2010). In addition to carbon catabolite repression, Hfq and Crc link key metabolic and virulence processes in *Pseudomonas* species. The two proteins together affect biofilm formation and motility (O'Toole *et al.*, 2000, Huang *et al.*, 2012; Zhang *et al.*, 2012; Pusic *et al.*, 2016), antibiotic susceptibility (Linares *et al.*, 2010; Heitzinger, 2016) and biosynthesis of the virulence factor pyocyanin (Sonnleitner *et al.*, 2003; Huang *et al.*, 2012). Recent ChiP-seq studies in *Pseudomonas* indicate that Hfq and Crc also have wider regulatory impact and can co-operate to bind many nascent transcripts co-transcriptionally (Kambara *et al.*, 2018). As such, these transcripts form a large pool of potential regulatory targets for Hfq and Crc, some of which have been studied in this chapter.

Crc and Hfq act together to capture a target mRNA and prevent its translation. The targets are then subjected to degradation, which might trigger disassembly of the Hfq-Crc-RNA complex (Sonnleitner & Bläsi, 2014). It was found that Hfq and Crc-mediated regulation of translational repression involves binding of both proteins to ribosome-binding sequences of target mRNAs (Figure 2). The best studied target mRNA is *amiE*, which encodes the enzyme Aliphatic amidase that removes the amide group of short-chain aliphatic amides, generating their corresponding organic acids. This chemical transformation enables *Pseudomonas* to utilise acetamide as a source of both carbon and nitrogen. However, when preferred carbon sources such as succinate are abundant, Aliphatic amidase is not needed, and the CCR control system will repress the unnecessary translation of *amiE* mRNA by sequestering the *amiE* ribosome binding site (RBS) with Hfq. The interaction occurs through the distal face of Hfq, which engages an A-rich segment called $amiE_{6ARN}$ (Sonnleitner & Bläsi, 2014). Strikingly, Crc was found to strengthen binding of A-rich target transcripts to the distal side of Hfq, while having no intrinsic RNA-binding activity on its own (Sonnleitner *et al.*, 2018; Milojevic *et al.*, 2013). The translational repression effect of Hfq and Crc on target mRNAs is countered by the non-coding RNA *CrcZ*, which is believed to sequester Hfq and Crc (Sonnleitner *et al.*, 2009; Sonnleitner & Bläsi, 2014) (Figure 2). *CrcZ* levels increase when preferred carbon sources are exhausted, and its expression is under control of the alternative sigma factor RpoN and the two-component system CbrA/B (Sonnleitner *et al.*, 2009). The signal responsible for CbrA/B activation remains unknown, but it is thought to be related to the cellular energy status (Valentini *et al.*, 2014).

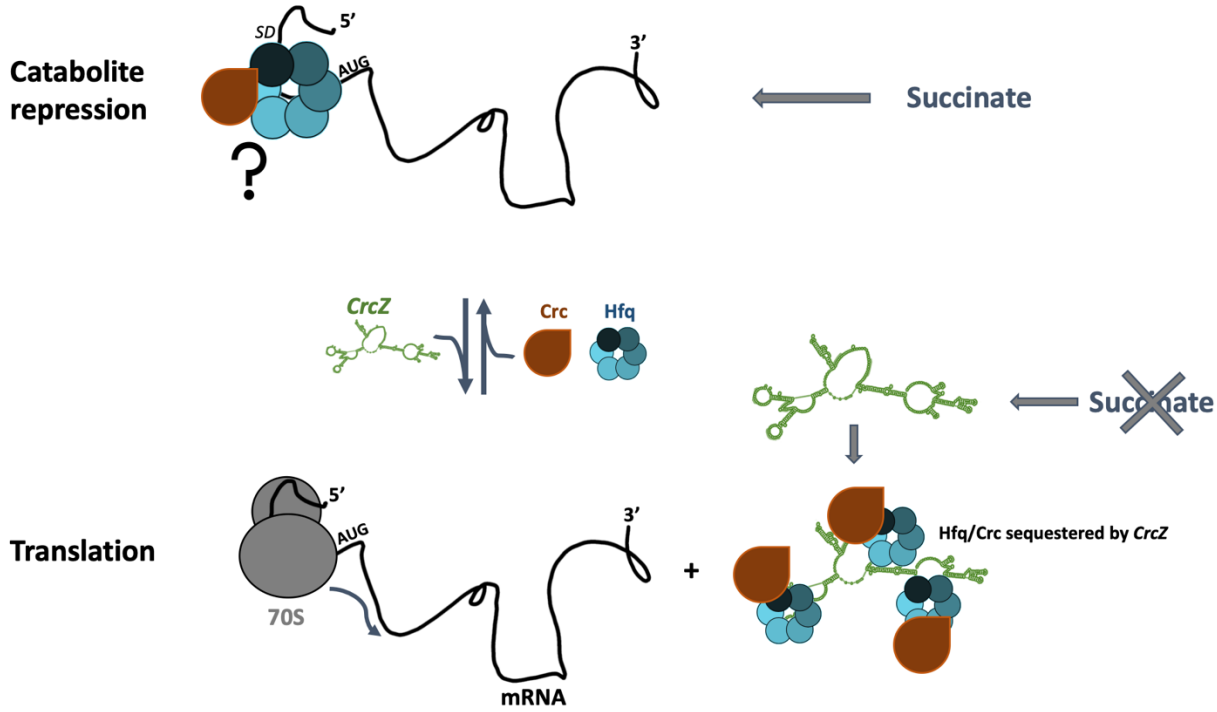


Figure 2: ***Catabolite repression in Pseudomonads.*** *During catabolite repression, e.g. when succinate levels are high, Hfq and Crc cooperatively bind A-rich sequences at the 5'-end of the target mRNA, amiE, and mask the ribosome binding site (top). When succinate levels are low, catabolite repression is alleviated by expression of the sRNA CrcZ, which sequesters Hfq and Crc from target RNAs (bottom).*

How Hfq binds to A-rich RNAs is understood structurally (Link *et al.*, 2009), but the co-operative role of Crc to prevent translation of a target RNA in this context has not been elucidated. Furthermore, it is unclear how the same two effector molecules can co-regulate more than 100 target RNAs co-transcriptionally with apparent sequence specificity. To gain insight into how *P. aeruginosa* Hfq cooperates with Crc in CCR-mediated translational repression of *amiE*, structures were solved of the complexes the two proteins form on the 18-residue long Hfq binding motif of *amiE* mRNA using cryo-electron microscopy (cryo-EM). The reconstructions revealed for the first time how presentation of RNA in a sequence-specific manner can support the association of Hfq and RNA with Crc. Mutagenesis of the interactions between e.g. Crc and *amiE* guided by the structural data corroborate their requirement for *in vivo* translational regulation. Molecular dynamics studies of these assemblies reveal a putative dynamic interface at the Hfq distal side that allows for rapid scanning of large RNA pools in the cell. Additional cryo-EM studies on a 105-residue long portion of the *amiE* 5'-end provide insights into the quaternary structure of the translation-repression assembly and reveal potential kinetic intermediates. Gel shift assays further show that similar assemblies form on other predicted target RNAs such as transcripts for *estA*, *acsA*, *rbsB* and *ptsX* (Kambara *et al.* 2018). Cryo-EM structures of Hfq-Crc assemblies on one of these target mRNAs, *rbsB,* provide a structural basis that explains how two effector molecules can co-regulate a large pool of mRNA targets in *Pseudomonads*. These findings expand the paradigm for *in vivo* action of Hfq through cooperation with the Crc helper protein and RNA to form effector assemblies that regulate target gene expression.

3 Results

3.1 Synthesis and purification of Hfq and Crc

The expression and purification of *Pseudomonas* (PAO1) Hfq and Crc is explained in detail in the materials and methods section (Sections 5.1 and 5.2). Figure 3 shows representative elution profiles and SDS PAGE gels of hexameric Hfq (56kDa) and monomeric Crc (28kDa) after a size exclusion chromatography step. Both samples are sufficiently homogeneous and pure for further experiments. A small shoulder was observed in the Hfq elution profile (Figure 3A, top) and SDS/heating were insufficient to fully denature the heat resistant Hfq hexamer, as is apparent from the SDS PAGE gel (Figure 3A, bottom, Hfq oligomers marked by *). The Crc elution profile reveals a slightly asymmetric peak and faint contaminant bands (marked by * in Figure 3B) are visible on the SDS PAGE gel. Since PAO1 Crc samples synthesized in *E. coli* have shown minor ribonuclease activity even after purification, these were later replaced with Crc samples synthesized and purified from *P. aeruginosa* (PAO1), which are free of ribonuclease contamination. These purified Crc samples were prepared by collaborators (Bläsi group, MFPL, University of Vienna). All RNA samples described below were synthesised chemically or prepared with *in vitro* transcription and subsequent gel-purification/electroelution by the Bläsi group.

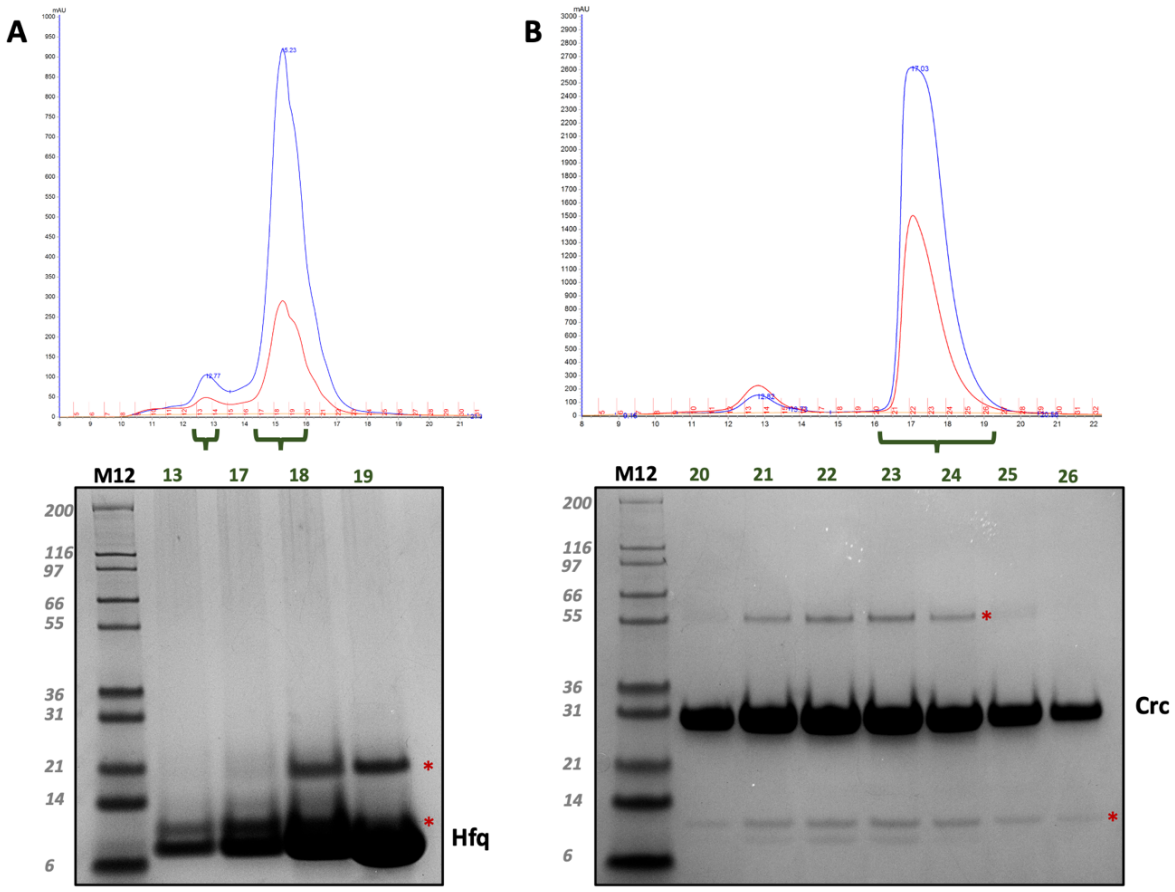


*Figure 3: **Elution profiles and SDS PAGE gels for Hfq (A) and Crc (B) after size exclusion chromatography**. Both Hfq and Crc are homogeneous and pure enough for further experiments. * highlights Hfq oligomer bands (A) or contaminants (B).*

3.2 Cryo-EM analysis of Hfq-Crc-*amiE*$_{6ARN}$: An ensemble of assemblies

Previous SEC-MALS and mass spectrometry studies have suggested that Hfq and Crc form a higher order translation-repression complex with *amiE* (Sonnleitner *et al.*, 2018). Purified Hfq and Crc proteins were mixed with a chemically synthesized 18-nucleotide Hfq-binding element from the translation initiation region of *amiE* mRNA (*amiE*$_{6ARN}$) (5'-AAA-AAU-AAC-AAC-AAG-AGG-3'). This sequence has a binding motif comprised of six repeats of an A-R-N pattern known to bind the distal face of Hfq (Sonnleitner & Bläsi, 2014). After mild chemical crosslinking with BS3 (see Materials and Methods section 5.4), the sample of Hfq-Crc/*amiE*$_{6ARN}$ was loaded on R1.2/1.3 Quantifoil grids decorated with a graphene oxide monolayer (see Materials and Methods section 5.5). Grid screening revealed well defined single particles on graphene oxide in regions with thicker vitreous ice, and datasets were collected on a Titan Krios equipped with a Falcon III detector (FEI). For these datasets all processing was performed in Relion 3.0 (Zivanov *et al.*, 2017). Selective 2D classification and subsequent 3D classification revealed three principal types of complexes corresponding to different stoichiometries of Hfq:Crc:*amiE*$_{6ARN}$ (Figure 4). These higher order assemblies, with Hfq:Crc:*amiE*$_{6ARN}$ compositions 2:2:2, 2:3:2 and 2:4:2, are in accord with published SEC-MALS and mass spectrometry findings that excluded a more simple 1:1:1 assembly (Sonnleitner *et al.*, 2018).

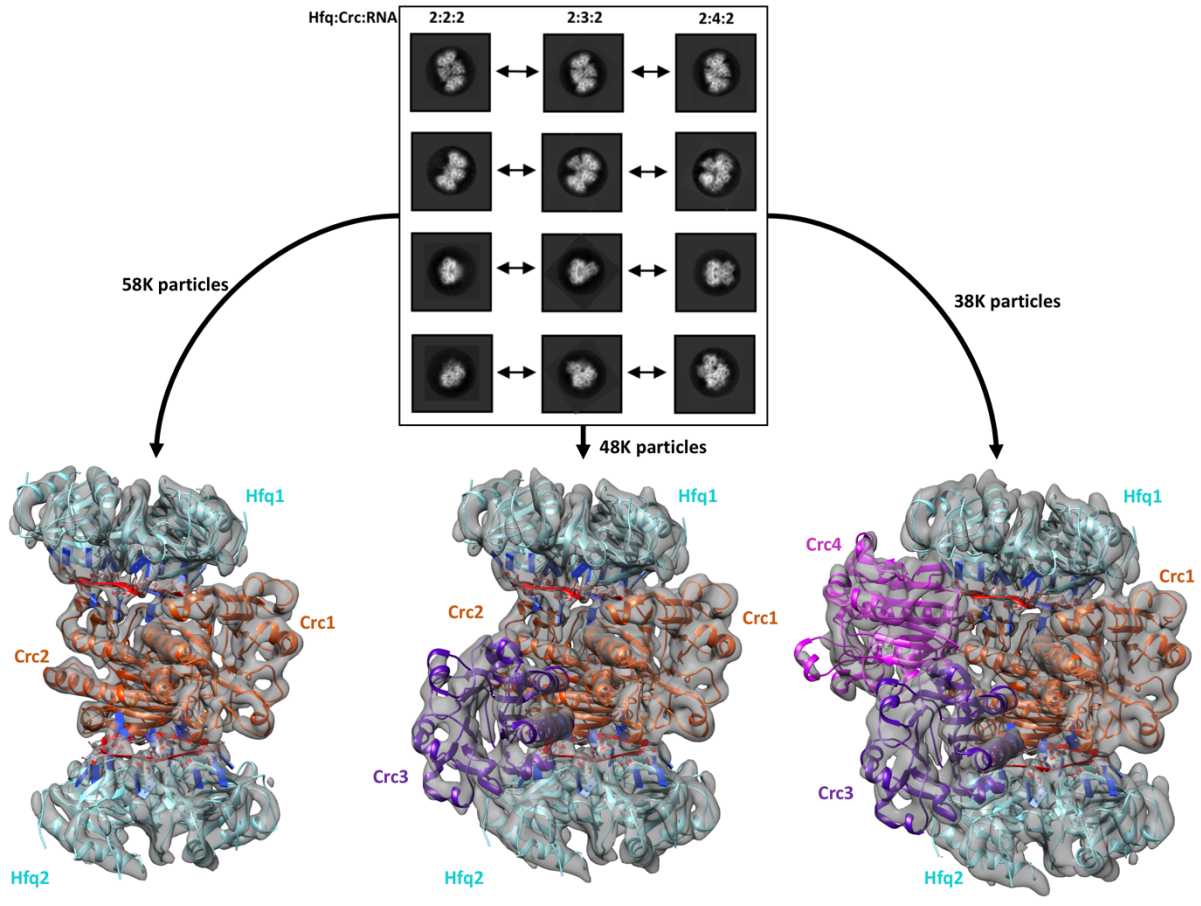


Figure 4: ***2D- and 3D classification of Hfq-Crc-RNA particles****. Three main classes of particles were observed after reference-free 2D classification (top), corresponding to Hfq:Crc:amiE*$_{6ARN}$ *assemblies with compositional stoichiometries of 2:2:2, 2:3:2 and 2:4:2. The amiE*$_{6ARN}$ *species (red) constitute the main interaction interface between Hfq and Crc, together forming the 2:2:2 core complex observable in all three models (bottom). Hfq hexamers are depicted in cyan, Crc in orange, purple or pink. All cryo-EM maps are low pass filtered to 6 Å for interpretability. Only a subset of the 2D classes are shown in the top panel.*

The individual assemblies appear to be conformationally uniform and well defined, and the refined maps for these assemblies are estimated to be 3.1 - 3.4 Å in resolution based on gold-standard Fourier shell correlations (GS-FSC) (Figure 5, Table 1). Crystal structures of Hfq and Crc fit well into the cryo-EM densities, apart from a few side chain rotations into alternative permitted rotamers (Figure 4 and Figure 6). All 18 nucleotides of *amiE*$_{6ARN}$ were resolved to near-atomic resolution in the three cryo-EM maps and will be referred to with numbers 1–18 from the 5´- to the 3´- end. To account for potential crosslinking artefacts, Cryo-EM analyses were subsequently performed with samples that had not been treated by crosslinking using the same grid preparation conditions optimised for the crosslinked specimens (data not shown). Data from these specimens revealed that the quaternary structure remained unchanged within the resolution limits of the data.

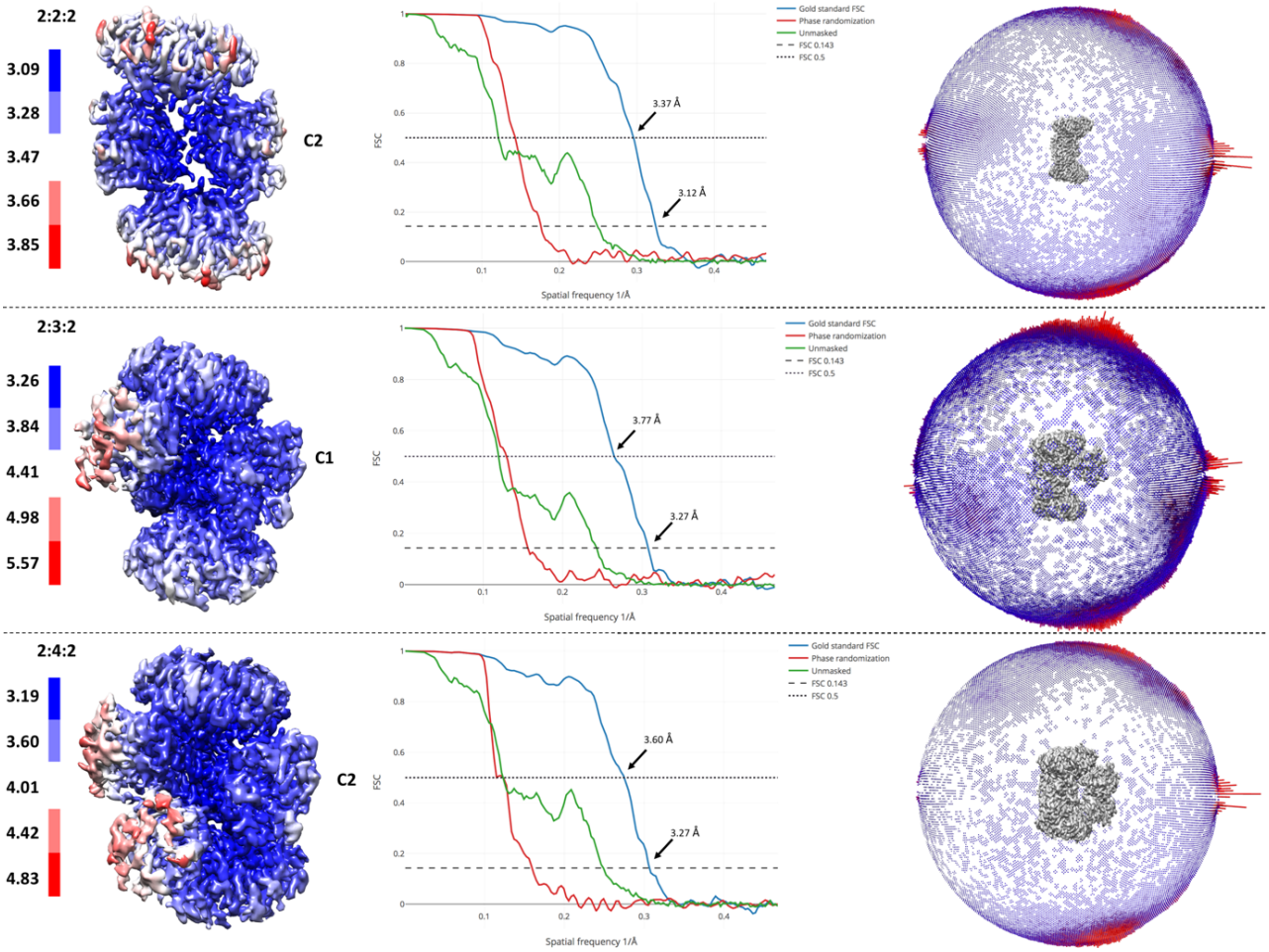


Figure 5: ***Global and local resolution estimates of the cryo-EM maps****. GS-FSC metrics were used for both local and global resolution estimates. Global FSC curves for all three reconstructions are depicted in the central panel. FSC 0.143 and 0.5 are annotated. In the right panel, angular distributions of the 2D images are presented as a spherical plot, with red bars representing more abundant projections in the images. C1, No symmetry; C2, C2 symmetry.*

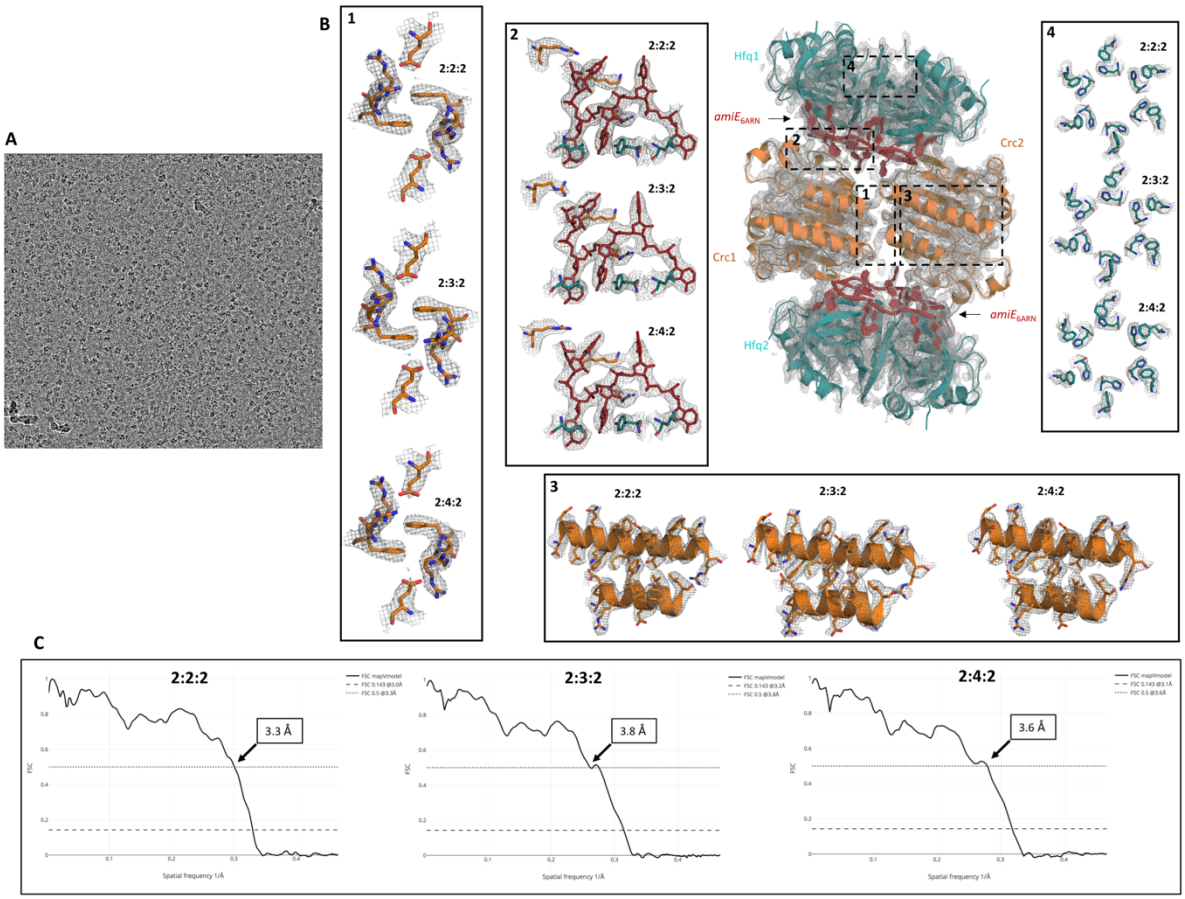


*Figure 6: **Quality of the cryo-EM images and reconstructions**. **(A)** Example of a raw micrograph at 3 µm defocus, lowpass filtered to 20 Å. **(B)** High resolution cryo-EM maps with refined atomic models for all complexes, showing the quality of the EM reconstructions. All maps were reconstructed independently of each other or a high-resolution reference structure and reveal well defined and highly reproducible densities for all side chains in the highlighted signature regions (insets 1-4). At the periphery, the map density is of good quality too, maintaining the six-fold symmetry of the Hfq protomers (inset 4). **(C)** Model versus map Fourier shell correlation curves (FSC) show a good correlation between the atomic models and the experimental cryo-EM maps. FSC 0.5 is annotated on the graph.*

3.2.1 The Brennan crown, a recurring RNA fold in Hfq-mediated regulation

A previous crystal structure of *E. coli* Hfq bound to a polyriboadenylate 18-mer solved by the Brennan group revealed how the RNA encircles the Hfq distal face via a repetitive tripartite binding scheme (Figure 7B) (Link *et al.*, 2009). In that structure, base triplets occupy inter-monomeric pockets on the Hfq surface and fold into a crown-like conformation. As the Brennan group first described this RNA binding mode, the fold will be referred to as the 'Brennan crown' in this and subsequent chapters. In the structures presented here, the RNA constitutes the binding interface between *P. aeruginosa* Hfq and Crc and adopts a nearly identical crown-like conformation on the Hfq distal face as seen in the *E. coli* Hfq/poly-A complex (Figure 7A and B). Since the cryo-EM maps were calculated without any reference to the structure from Link *et al.* (2009), the models presented here not only are a strong validation of the broader occurrence of such regulatory RNA folds, but also for the first time show their role in the setting of a biologically relevant ribonucleoprotein assembly.

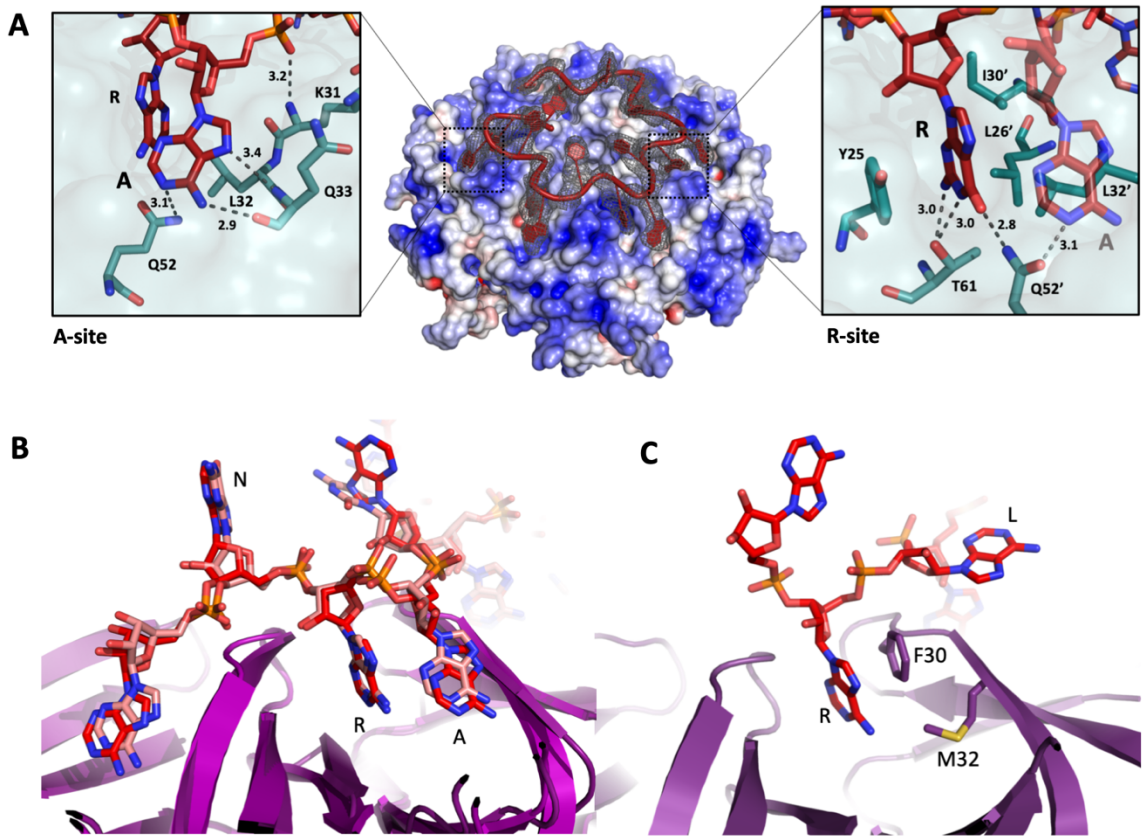


Figure 7: ***The "Brennan crown" in the Hfq/amiE$_{6ARN}$ RNA complex****. **(A)** Six RNA triplets are partially embedded in six binding pockets on the Hfq distal side, forming a weaving, crown-like pattern. The RNA entry/exit site has no discriminatory preferences and is therefore referred to as the 'N' site, whereas the A and R sites are occupied by adenine and a purine, respectively. The RNA crown is modelled in red and the cryo-EM density for amiE$_{6ARN}$ is depicted as a grey mesh. The RNA folds into a star-shaped conformation guided by positively charged protuberances on the Hfq distal surface (blue). Left panel: adenosine specificity site. Right panel: Purine specificity site. amiE nucleotide carbon atoms are depicted in red, Hfq carbon atoms are in green. Hydrogen bond lengths are in Å (dashed lines). **(B)** Overlay of the tripartite pocket between E. coli and P. aeruginosa Hfq and poly-A$_{18}$ and amiE$_{6ARN}$ RNA, respectively. The crystal structure of E. coli Hfq in complex with poly-A$_{18}$ (red) is strikingly congruent with the interactions between P. aeruginosa Hfq and amiE$_{6ARN}$ RNA (light pink) in the complex with Crc. **(C)** The A-R-N motif is not supported by Hfq of the Gram-positive bacteria Staphylococcus aureus and Bacillus subtilis, which instead use a R-L (purine-linker) motif (Horstmann et al., 2012). This can be rationalised by the pattern of amino acids that occlude the corresponding A site.*

Like its *E. coli* homologue, the *P. aeruginosa* Hfq distal side contains six tripartite binding pockets that can bind a total of 18 nucleotides. In addition, six positively charged protuberances on the same distal face guide the star-shaped RNA fold, where the phosphate backbone weaves in between these (Figure 7A, middle panel). Each tripartite pocket consists of an adenosine specificity site (A), a purine nucleotide specificity site (R), and a non-discriminatory RNA entrance/exit site (N) (Figure 7A and 6B). The A- and R-sites constitute a structural preference for RNA stretches with $(ARN)_n$ repeats. All bases in the A-sites of the *amiE*$_{6ARN}$ substrate are organized identically to the corresponding A-sites in the *E. coli* orthologue (Link *et al., 2009*), and are depicted in Figure 7A. Importantly, the interactions between adenosine and the carbonyl of Q33 and between adenine and the backbone amide of

Q33 confer pocket-specificity for the A-site as they are not compatible with a guanine base, which would form a repulsive contact via its O6 and peptide carbonyl (Figure 7A, left panel).

Two neighbouring monomers make up the purine (R) specificity pocket on the Hfq distal side. Unlike the A-site, which is on the Hfq surface, the R-site is tucked in and therefore less exposed (Figure 7A, right panel). Notably, Gln52 forms a physical link between the occupied A- and R-sites. The structures presented here were solved with the authentic *amiE* Hfq recognition site, whereas previous structures were obtained with A_{18} RNA (Link *et al.*, 2009). In those earlier structures all R-site were occupied by adenine residues. In the structures presented here, Thr61 forms a double hydrogen bond with a guanine base in the R-site, which has not been observed before (Figure 7A).

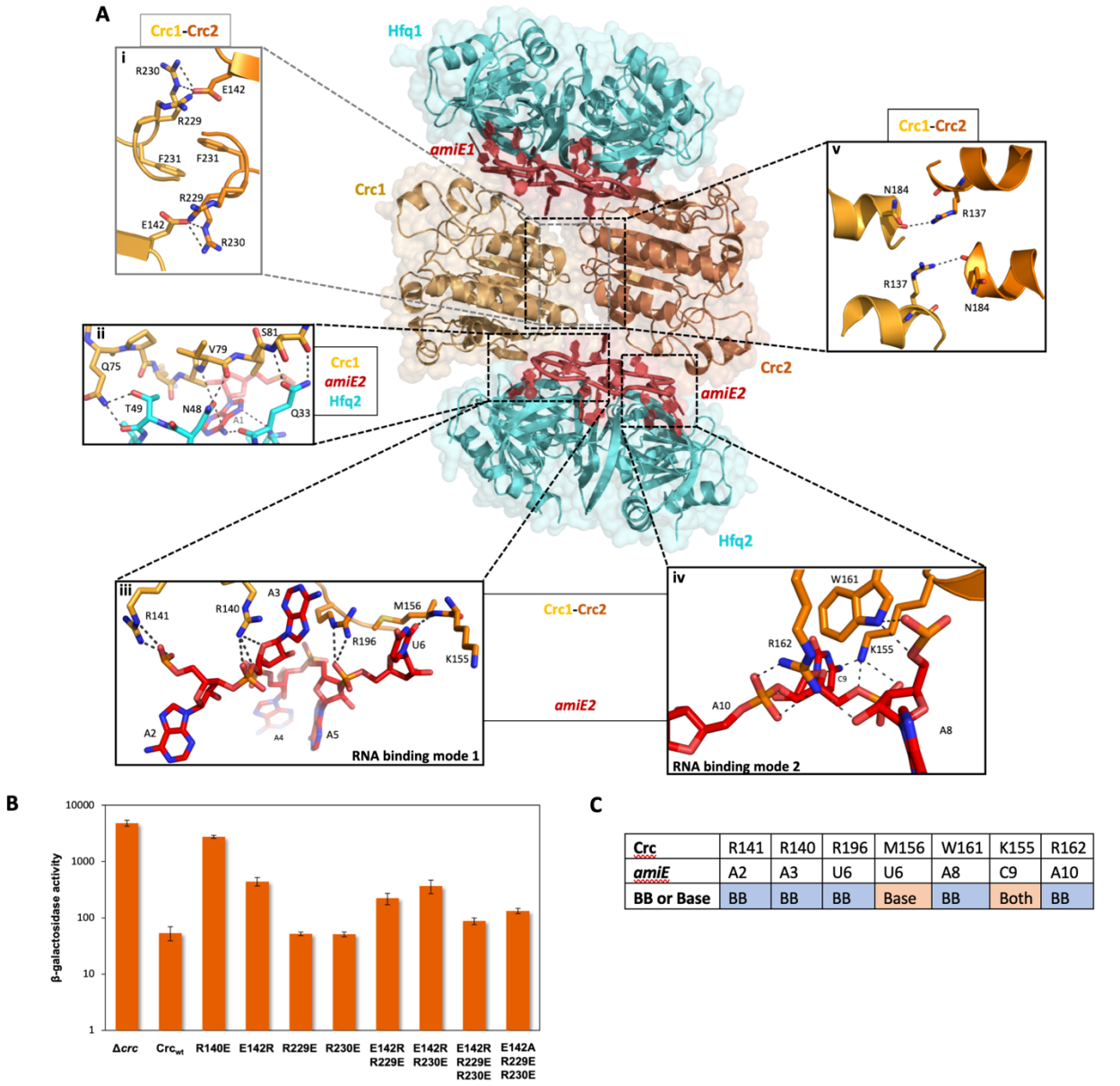
The Hfq-*amiE*$_{6ARN}$ interactions are not influenced by the bases in the entry/exit (N) sites, and one would expect the latter to participate in interactions with effector molecules in the assembly. Indeed, N-site bases interact with Crc in the cryo-EM models, as will be described below. Strikingly, the ARN-preference of the Hfq distal side is not shared in Gram-positive bacteria. In *Staphylococcus aureus* and *Bacillus subtilis*, for example, Phe30 in the corresponding A-site sterically hinders the RNA backbone and A-site base from entering, yet stabilises purines in the R-site together with M32 via stacking interactions (Figure 7C) (Horstmann *et al.*, 2012). Hfq from these Gram-positive bacteria thus prefers an R-L (purine, Linker) motif.

3.2.2 Interactions in Hfq-Crc-*amiE*$_{6ARN}$ core complex

The core complex consists of two Hfq molecules each with an *amiE*$_{6ARN}$ motif bound, which together sandwich two copies of Crc (Figure 8A). Given the short length of the *amiE*$_{6ARN}$ fragment, only Crc-RNA and Crc-Hfq interactions will be discussed here, as the quaternary structure of the assembly is unlikely to fully capture the subunit organisation that occurs on a long, continuous RNA that can span all potential binding sites on the proteins. The true quaternary structure on such an RNA will be presented in section 3.4.

Both Crc molecules interact with exposed *amiE*$_{6ARN}$ on the Hfq surface as well as with the Hfq distal rim (Figure 8A). Mainly the phosphate backbone and exposed ribose rings of the RNA contact the Crc molecules (Figure 8A and C). Two Crc monomers form an antiparallel dimer, with each monomer having two modes of RNA binding. In one mode, Crc Arg140 and Arg141 form polar contacts with the phosphodiester backbone of *amiE*$_{6ARN}$. The purine-base of the A3 nucleotide at an Entry/Exit site of *amiE*$_{6ARN}$ is sandwiched by Crc Arg140 and Arg196 (Figure 8A, inset iii). Lastly, *amiE*$_{6ARN}$ U6 forms a hydrogen bond with the Crc Met156 amide, and Crc Arg196 forms hydrogen bonds with the U6-*amiE*$_{6ARN}$ backbone. For the second mode of RNA binding, constituted by a second Crc surface, Crc Lys155 makes a hydrogen bond with the phosphate group of C9 and the ribose hydroxyl group. Crc Trp161 and Arg162 form additional hydrogen bonds with the *amiE*$_{6ARN}$ backbone (Figure 8A, inset iv, and C). Only a few direct Hfq-Crc contacts can be observed in the cryo-EM structures. In short, Crc Gln75, Val79 and Ser81 form a small interaction β-strand on each of both Crc components in the core complex, contacting Thr49, Asn48 and Gln33, which are all presented on a short loop of a single Hfq protomer (Figure 8A, inset ii). In summary, the highly

organised interactions in the core complex show how Hfq presents *amiE*$_{6ARN}$ to form a molecular interface for the RNA-mediated interactions between Hfq and Crc (Figure 8A).



Crc	R141	R140	R196	M156	W161	K155	R162
amiE	A2	A3	U6	U6	A8	C9	A10
BB or Base	BB	BB	BB	Base	BB	Both	BB

Figure 8: ***Model of the 2:2:2 Hfq-Crc-RNA core complex and validation of interactions. (A)*** *Atomic model of the core assembly. Hfq hexamers sandwich the antiparallel Crc dimer and present the amiE$_{6ARN}$ RNA to form two different interfaces with the Crc protomers. The Crc molecules form polar contacts with mainly the exposed ribose rings and backbone phosphate groups (insets iii and iv). Only a small binding interface on each Crc monomer is involved in direct contacts with a Hfq protomer (inset ii). Crc protomers are in yellow and orange, amiE$_{6ARN}$ RNA is in red, Hfq hexamers are in cyan.* ***(B)*** *Translational regulation of an amiE::lacZ reporter gene by Crc variants, as designed based on the cryo-EM model. P. aeruginosa strain PAO1Δcrc(pME9655) harbouring plasmids pME4510 (vector control), pME4510crc$_{Flag}$(Crc$_{wt}$) or derivatives thereof encoding the respective mutant proteins were used. The beta-galactosidase values conferred by the translational amiE::lacZ fusion encoded by plasmid pME9655 in the respective strains are indicated. The results represent data from two independent experiments and are shown as mean and standard deviation.* ***(C)*** *Table of direct Crc-amiE$_{6ARN}$ contacts in the 2:2:2 assembly. BB: Crc contacts with the RNA backbone; Base: Crc contacts with RNA bases.*

3.2.3 Further masking of *amiE* by Crc, the 2:3:2 and 2:4:2 complexes

The interaction of Crc with RNA in the core complex leaves approximately half of the accessible surface of *amiE*$_{6ARN}$ exposed. Additional Crc units can be recruited through interactions with the exposed portion of the RNA, forming 2:3:2 and 2:4:2 complexes (Figure 4 and Figure 9). Although the core complex is rigid, the additional Crc molecules (Crc3 and Crc4) have few contacts with the core complex and each other, and they are likely to be comparatively mobile as evidenced by their somewhat lower local resolution in the cryo-EM maps (Figure 5). Both Crc3 and Crc4 interact with the two amiE$_{6ARN}$ motifs via the first Crc-RNA binding mode (Figure 9A, inset iii). Crc3/4 K139 makes a hydrogen bond with the *amiE*$_{6ARN}$ C12 phosphate group, Arg138 interacts with the ribose hydroxyl group of C9 and K135 forms a hydrogen bond with the A11 phosphate group, all of which were left exposed in the 2:2:2 core complex. Finally, the backbone amino group of Arg140 engages in a hydrogen bond with the C12 base. Direct contacts between Crc3/4 and Hfq in the 2:4:2 cryo-EM model are constituted by the same Crc β-strand and exposed loop of a single Hfq protomer (Figure 9A, inset ii).

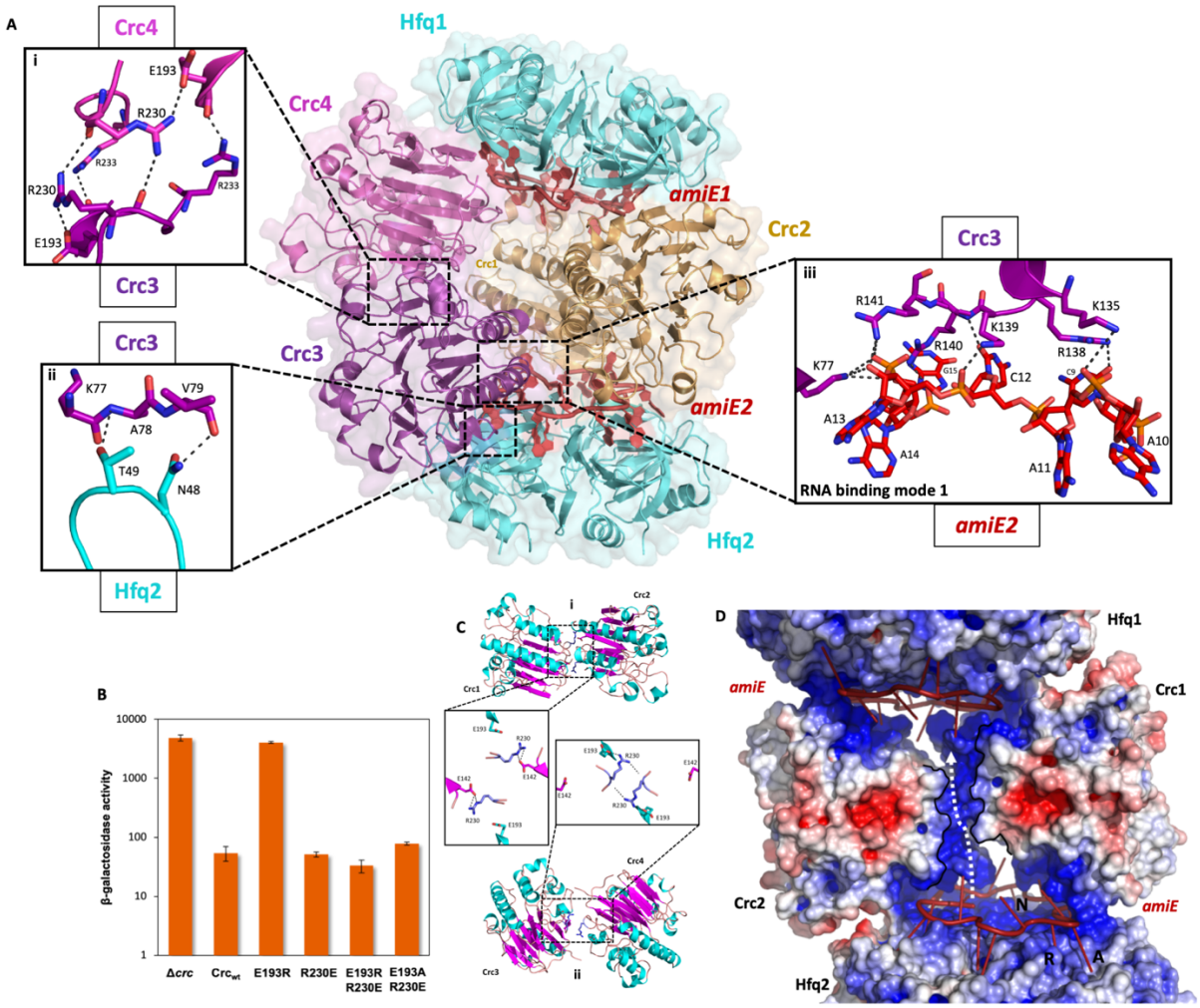


Figure 9: ***Model for the 2:4:2 Hfq-Crc-RNA complex. (A)*** *Atomic model of the 2:4:2 Hfq-Crc-RNA complex. The insets show the Crc-Crc (i), Hfq-Crc (ii) and Crc-RNA (iii) interactions not present in the 2:2:2 complex. Only one of two Crc RNA-binding patches is presented to amiE*$_{6ARN}$ *in the Crc3-Crc4 dimer (inset iii). A small interface is formed between Crc3/Crc4 and Hfq. The core Crc dimer is in yellow, amiE*$_{6ARN}$ *RNA is in red, Hfq hexamers are in cyan, the protomers of the extra Crc dimer are magenta and purple.* ***(B)*** *Translational regulation of an amiE::lacZ reporter gene by Crc variants. The results represent data from two*

independent experiments and are shown as the mean, with standard deviation bars. ***(C)*** *Two distinct Crc dimers are observed in the three complexes solved by cryo-EM. i: The self-complementary interaction of the 2:2:2 core complex. ii: an alternative dimer is formed in the 2:4:2 complex, showing an open configuration (bottom).* ***(D)*** *An electropositive half-channel runs along the core dimer interface (Crc1-Crc2) which could potentially serve as a conduit for RNA (dotted white arrow). The A, R, and N RNA interaction sites on the Hfq distal side are annotated.*

Interestingly, there are two Crc dimer conformations in the 2:4:2 complex. A first Crc dimer is formed by Crc1-Crc2 in the core assembly and is identical to a Crc dimer observed in the crystal lattice (pdb-ID 4jg3) (Figure 9C). A second Crc dimer is formed by Crc3-Crc4 and adopts a more open conformation, although preserving the dimer interface (Figure 9C). These observations are a first hint that Crc dimerization is polymorphic, which will be addressed extensively in section 3.4. A basic, surface-exposed half-channel runs through the Crc1-Crc2 dimer interface (Figure 9D). Speculatively, longer RNA species, such as the full length *amiE* transcript, could run through such a half-channel and interconnect all components of the assembly into a highly organised complex. In section 3.4.2, structural studies on a longer fragment of the 5'-end of *amiE* will show that this hypothesis is not too far off.

3.2.4 *In vivo* validation of the Hfq-Crc-*amiE$_{6ARN}$ interactions*

To probe the relevance of the Hfq-Crc-*amiE$_{6ARN}$* interactions, colleagues at the Bläsi group (MFPL, University of Vienna) tested the effects of specific Crc mutants on Hfq-Crc-mediated repression of an *amiE::lacZ* translational reporter gene *in vivo*. Firstly, the interaction between Crc and the presented *amiE$_{6ARN}$* fold was tested. R140 (Crc) interacts with *amiE$_{6ARN}$* in the cryo-EM models (Figure 8A, inset iii, Figure 9A, inset iii). A Crc$_{R140E}$ mutant was deficient in translational repression of the *amiE::lacZ* reporter gene, with levels similar to those observed for a Crc deletion strain (Figure 8B). In addition, Crc$_{R140E}$ did not form higher order assemblies with Hfq and *amiE$_{6ARN}$* in co-immunoprecipitation experiments (data not shown, see Pei *et al.* 2019).

Next, the relevance of the Crc dimer interfaces was tested. Reporter assays with different mutants of the Crc1-2 dimerization patches show only a mild reduction in translation-repression, which can be compensated by counter mutations (Figure 8A and B, discussed in detail in Pei *et al.*, 2019). The Cryo-EM models presented in section 3.4 indicate that the Crc1-2 dimer is not formed in the true quaternary structure of the *amiE* translation repression assembly. Notably, the Arg233-Glu193 interaction in the Crc3-4 dimer, on the other hand, is key for adequate translation-repression (Figure 9A, inset i). A Crc$_{E193R}$ substitution was sufficient to fully abrogate translational repression of *amiE::lacZ* (Figure 9B). According to the cryo-EM model, the deleterious Crc$_{E193R}$ substitution should be compensated by a Crc$_{R230E}$ mutation and re-establish the interaction. The *amiE::lacZ* reporter assay showed that Crc$_{E193R, R230E}$ and Crc$_{E193A, R230E}$ variants restored translational repression. Since the Crc variants are all expressed at comparable levels to the wild type Crc (data not shown), the reporter assays presented here show that the interactions present in the Hfq-Crc-*amiE$_{6ARN}$* models are relevant *in vivo*.

3.3 Intrinsic molecular scanning opportunities at the Hfq distal side

The cryo-EM models described in section 3.2 provide insights into how a specific target-mRNA (in this case *amiE*) can be bound and presented by an RNA chaperone to trigger assembly of an intricate effector complex that represses translation. The results, however, raise questions regarding proofreading mechanisms in the Hfq-Crc-*amiE* assembly process. One question is related to how Hfq, which binds many RNAs in the cell, can recognise or scan for ARN repeats in a given sequence. To explore this question, molecular dynamics (MD) simulations were carried out with collaborators Prof. Jiri Sponer and Dr. Miroslav Krepl (CEITEC, Brno, Czech Republic) (Krepl *et al.*, in press).

3.3.1 Dynamic recognition at the Hfq distal side: *syn*/*anti* flipping of the A-site bases

Hfq is perhaps one of the most pleiotropic RNA chaperones in bacteria, binding many different sRNA-, mRNA- and tRNA species (Vogel & Luisi, 2011; Melamed *et al.*, 2016, 2020). A puzzling question is how Hfq recognises and distinguishes between different RNA molecules. Finding targets likely involves a mechanism of scanning an RNA molecule on the Hfq distal side until ARN-rich regions are encountered. If the RNA molecule harbours such a region, assembly formation with Crc partners likely follows. If not, the RNA dissociates from the Hfq distal side and a different RNA is scanned. A recent study suggests that protein-RNA recognition could involve so called 'dynamic recognition', whereby protein-RNA interactions occur via competing local conformational sub-states of the interface (Ripin *et al.* 2019), as distinct to a 'stiff geometry' binding. Dynamic recognition could allow for high binding turnover *in vivo* to quickly probe large pools of cellular RNAs, and at the same time maintain their highly specific recognition of RNA targets by proteins. It is likely that Hfq utilises some sort of dynamic recognition mechanism, given its disproportionally high cellular RNA turnover relative to its low-nanomolar affinity for target RNAs (Fender *et al.*, 2010).

MD simulations of the Hfq-Crc-*amiE*$_{6ARN}$ assembly and sub-assemblies indeed reveal how extensive dynamics occur at the Hfq/RNA interface. In the cryo-EM structures, all A- and R-site bases are in the *anti*-conformation (Figure 7A) when Crc is bound, and their binding pockets are described in detail in section 3.2.1. Importantly, there were no signs of structural instability during the simulations, indicating that the starting models were of decent quality and that the force field was performing well. Strikingly, a dynamic equilibrium was observed in the absence of Crc molecules (i.e. for Hfq-RNA intermediates), where the *amiE*$_{6ARN}$ A-site bases flip between *syn*- and *anti*-conformations (Figure 10). Several back-and-forth *syn/anti* flips took place within the simulation timescale, with transitions occurring on the order of hundreds of nanoseconds. The *syn*- and *anti*- conformers of the adenine base both engage the exact same amino acids in the A-site pocket of the Hfq-distal side, albeit with different H-bond donor- and acceptor groups (Figure 10). The simulations suggest that in the absence of Crc components, the A-site bases occupy a *syn*-conformation more often than the *anti*-conformation. Local perturbations in the Hfq/RNA interface, constituted by *syn/anti* base flipping, could allow for other RNA molecules to more effectively compete with the engaged RNA. It should be noted that the base-flipping process can occur on one ARN repeat at a time, which may allow stepwise RNA release and exchange as proposed previously (Fender *et al.*, 2010).

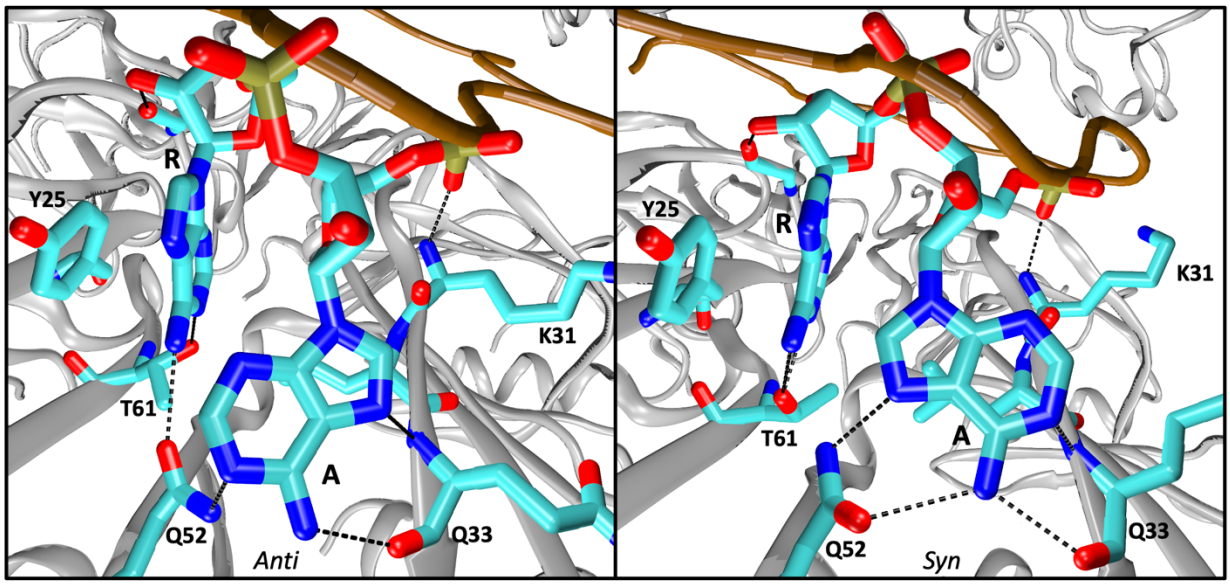


Figure 10: ***Anti/syn flipping in the Hfq A-site pocket.*** *The adenine base can flip from anti (left panel) to syn (right panel) in the A-site pocket on the Hfq distal side, while interacting with the same Hfq residues. In syn, the adenine N7 atom replaces the N1 atom as a hydrogen bond acceptor, while a hydrogen atom on the N6 amino group now serves as a hydrogen bond donor (right panel). Hfq secondary structure is in grey, RNA backbone is in brown, side chains and bases are in cyan.*

The MD simulations indicate that Crc binding to the *amiE*$_{6ARN}$ interface lowers the overall atomic fluctuations of the RNA backbone. This likely increases the barrier for *syn/anti* flips by sterically obstructing these in the Hfq-Crc-*amiE*$_{6ARN}$ complex, which results in a nearly full suppression of *syn/anti* flips in the MD simulations (within the timescales measured, i.e. 10 ms). In addition, A-site nucleotides A_1 and A_{13} of *amiE*$_{6ARN}$ form specific H-bonds with Crc K77, almost irreversibly stabilizing them in *anti* upon full assembly of the Hfq-Crc complex (not shown). Apart from A1 and A13, there are no apparent specific interactions that promote the experimentally observed *anti*-conformation of the A-site bases. However, in a system where all the A-site bases were manually flipped into *syn a priori*, the presence of Crc still allowed several A_A bases to return to the *anti*-conformation within the simulation timescale, but not back to the *syn*-conformation. This simulation indicates that Crc actively promotes the *anti*-conformation for the A-site nucleobase rather than merely suppressing *syn*/*anti* transitions. The structural basis for this could be non-specific interactions between Crc and the backbone phosphate groups immediately downstream of the A-site nucleotides. These interactions alter the dihedral angles observed for this backbone suite and promote backbone dihedral angles associated with the *anti*-conformation of A-site nucleobases. In summary, the observations above provide a rationale for how the Crc can prolong the RNA binding to Hfq by suppressing base flipping in the A-sites.

3.4 Hfq and Crc as global metabolic and virulence regulators in *Pseudomonas*

Extensive research has been done on Hfq-Crc mediated gene regulation in *Pseudomonas* and many studies indicate that the system is a critical node in diverse regulatory pathways in the cell. A key study by Kambara *et al.* (2018) defined a pool of more than 100 mRNA targets of Hfq and Crc via ChiP-seq experiments. Their findings show that Hfq engages mRNA targets as they spool from the RNA polymerase, to subsequently present them to Crc, and that Hfq and Crc act on these transcripts in a highly cooperative and interdependent manner. As such, Hfq and Crc are likely to be prevalent co-transcriptional regulators in *Pseudomonads*, controlling translation of many nascent target mRNAs. This means that the structures presented above are only the 'tip of the iceberg', and that similar assemblies can form on different mRNA targets. How Hfq-Crc complexes can assemble on different mRNAs, i.e. how Hfq presents different RNA targets with specificity and how Crc molecules accommodate this degree of variability to form translation-repression assemblies, will be discussed in this section.

3.4.1 Expanding the Hfq-Crc target pool

Five putative target RNAs were selected from the predicted mRNA target pool to further explore their association with Hfq and Crc (Kambara *et al.,* 2018). Four of these are involved in transport or utilisation of carbon sources: *estA,* encoding an esterase, *acsA*, encoding acetyl-coA synthetase, *rbsB*, encoding a putative ribose transporter and the previously studied *amiE*, encoding aliphatic amidase (Kambara *et al.*, 2018). The fifth target is *ptxS*, which encodes a post-transcriptional regulator that controls exotoxin A production (Swanson *et al.*, 1999), and thus coordinates virulence.

Rather than isolating short putative ARN repeats in these transcripts, a longer stretch surrounding the translation start codon was generated for each transcript via *in vitro* transcription (Bläsi group, MFPL, University of Vienna). The shortest RNA construct (*acsA*) was 70 nucleotides long, whereas the longest (*estA*) was 119 nucleotides long. This design allowed to study whether multiple, potentially degenerate ARN repeats on a single transcript could participate in a translationally repressive Hfq-Crc assembly (Figure 11). For *amiE*, a 105-mer was used that also contains the $amiE_{6ARN}$ motif described in the Cryo-EM studies described above. To explore whether these five transcripts participate in Hfq-Crc assemblies and to probe assembly stoichiometries, electrophoretic mobility shift assays (EMSA) were carried out (Figure 12). To set up the binding reactions, purified RNA constructs were mixed with increasing quantities of Hfq in the absence and presence of an excess of Crc before running the gel electrophoresis experiments.

Figure 11: ***Sequences of selected target RNAs.*** *ARN motifs are highlighted in yellow. AUG start codons are in bold. For amiE$_{105}$ a secondary, degenerate ARN repeat region is annotated.*

All RNA constructs were able to bind Hfq in the EMSA experiments (Figure 12, lanes 1-5). Interestingly, higher order Hfq-RNA assemblies are formed with increasing Hfq concentrations (Figure 12, green arrows), which indicates that multiple Hfq molecules can bind a single RNA target. Lane 6 of the gels in Figure 12 shows that Crc alone is not able to bind RNA. Upon addition of Hfq, however, Crc participates in higher order assemblies with the RNA target and Hfq (Figure 12, lanes 7-10). Here too the gels indicate there is a gradual assembly build-up for increasing Hfq concentrations. It appears that for all five RNA targets, a molar excess of 3-4 Hfq molecules per RNA molecule results in the highest order assembly (Figure 12, bands highlighted with *). As explained in section 3.1, Crc produced in *E. coli* (BL21 DE3) shows modest signs of ribonuclease contamination, which is apparent from the accumulation of small RNA species at the bottom of lane 6. Interestingly, for all RNAs tested (except for *estA*), addition of Hfq protects the RNA from degradation by the ribonuclease contaminant. Nevertheless, for all future studies Crc purified from PAO1 was used, without apparent ribonuclease contamination (Bläsi group, MFPL, University of Vienna). The EMSA results show that Hfq-Crc complexes can assemble on all of the predicted RNA targets tested here and suggest that multiple Hfq molecules can engage an RNA target.

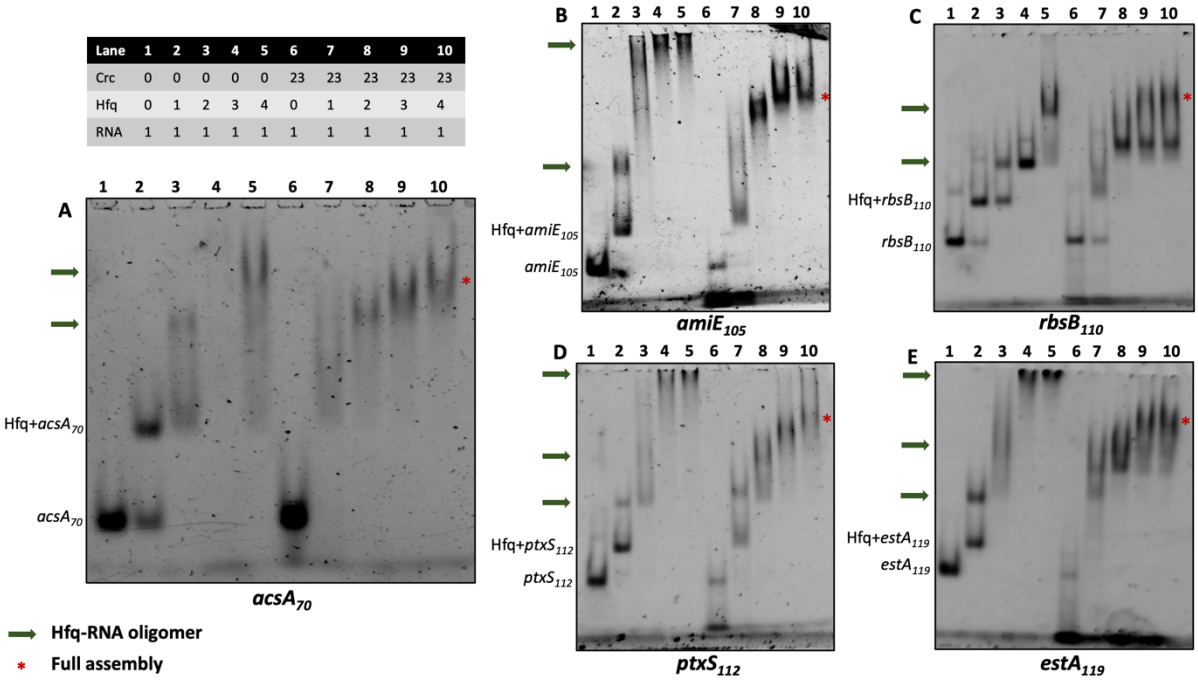


Figure 12: ***Electrophoretic mobility shift assays (EMSA) for predicted RNA targets.*** *Hfq and Crc can form higher order assemblies on all RNA targets tested here, acsA (A), amiE (B), rbsB (C), ptxS (D) and estA (E). For all RNAs tested, multiple Hfq molecules engage a single RNA molecule, both in the presence and absence of Crc. The table shows relative stoichiometries in the samples (RNA is at 200nM). Protein components were mixed first, after which the RNA was added. Green arrows highlight Hfq-RNA oligomers, * highlights the highest order complexes for the stoichiometry range tested.*

3.4.2 Cryo-EM analysis of the Hfq-Crc-*amiE*$_{105}$ assembly

The structures presented in section 3.2 contain only the *amiE*$_{6ARN}$ fragment of the *amiE* gene. Extrapolating from these structures, there are two possible scenarios for assembly. In one scenario, Hfq and Crc can trap two full *amiE* transcripts, by binding their respective ARN-triplet repeats. This would mean that two Hfq-*amiE* sub-assemblies have to come together in the cell to bind Crc. In a second scenario, however, a single *amiE* molecule threads through the complex and is presented by multiple Hfq molecules to several Crc protomers. For the latter, the Hfq hexamers in the assembly would accommodate one prefect ARN triplet repeat region (*amiE*$_{6ARN}$) and additional shorter/degenerate ARN repeat regions in the *amiE* sequence. To elucidate how many Hfq, Crc and *amiE* molecules participate in a translation-repression assembly, cryo-EM grids were prepared of Hfq-Crc-*amiE*$_{105}$. As described above, the RNA in these samples is a 105-residues long fragment of the *amiE* 5' region, containing multiple putative yet degenerate ARN repeat regions on top of the perfect 6xARN repeat (Figure 11). When screening the cryo-EM grids, many different intermediate assemblies were observed and potential insights from these intermediates will be further discussed in section 3.4.4. Only the final assembly, which was relatively abundant when mixing Hfq, Crc and *amiE*$_{105}$ at a molar ratio of 3Hfq:1amiE$_{105}$:11Crc, will be presented in this section.

In total, two Hfq-Crc-*amiE*$_{105}$ cryo-EM datasets were combined, both collected on a Titan Krios (FEI) equipped with a K3 detector (Gatan) (BioCem, Department of Biochemistry, University of Cambridge). A representative image at

2.5 µm defocus is shown in Figure 13A. For both datasets motion correction, CTF estimation and automated particle picking were carried out in Warp (Tegunov & Cramer, 2019) and 2D/3D alignments were done with cryoSPARC (Punjani *et al.*, 2017, Punjani *et al.*, 2019). After several rounds of 2D and 3D classification a consensus map was generated containing clear density for three Hfq hexamers and four Crc proteins, together forming a compact assembly on *amiE*$_{105}$ (Figure 13B and C). In other words, a single *amiE*$_{105}$ is presented to Crc by three Hfq chaperones. In the assembly, the two outer Hfq hexamers have their distal side facing each other (Figure 13C, Hfq1 and Hfq3), sandwiching the RNA substrate and the Crc proteins bound to it, similar to the structures described in section 3.2. The middle Hfq (Figure 13C, Hfq2) coordinates the RNA substrate via its distal side, proximal side and the rim side. In addition, an A-form hairpin RNA-fold is recognisable at the proximal side of the middle Hfq (Hfq2) (Figure 13C). After 3D refinement and global/local CTF refinements a map was reconstructed at 3.6 Å resolution (GS-FSC) with local resolution estimates ranging from 3.4 Å in the central map regions to 9 Å at the periphery of the cryo-EM density (Figure 14A and B). At these resolutions Hfq and Crc structures could be readily docked and refined and large stretches of the *amiE*$_{105}$ sequence could be traced and built *de novo* (Figure 14C) (Table 1). A detailed description of the Hfq-Crc-*amiE*$_{105}$ model will be given in section 3.4.3.

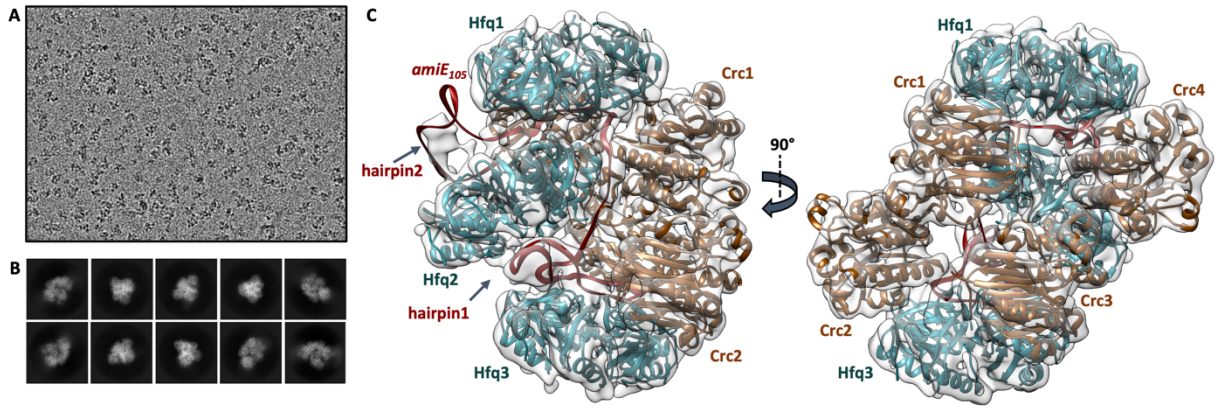


Figure 13: ***Cryo-EM reconstruction of the Hfq-Crc-amiE$_{105}$ assembly. (A)*** *Representative image at 2.5 µm defocus, lowpass filtered to 20 Å.* ***(B)*** *Selection of 2D class averages.* ***(C)*** *Three Hfq hexamers and four Crc molecules engage amiE$_{105}$ to form a translation-repression complex. The cryo-EM map is lowpass filtered to 7 Å for interpretability.*

Interestingly, when plotting the distribution of particle projections that contributed to the refined map, there is a noticeable difference in the orientation pattern for each of the two datasets (Figure 14D). The two datasets were collected from different cryo-EM grids, both coated with a layer of graphene oxide (GO) (see Materials and Methods section 5.5). For the first dataset (Figure 14D, top), the GO-coated grids were left at room temperature for 24 hrs before applying the Hfq-Crc-*amiE*$_{105}$ sample. For the second dataset (Figure 14D, bottom), the GO-coated grids were left for 48 hrs prior to sample application. For the latter grid it is clear that particle orientations are more diverse. Although more tests are required to confirm this observation, it is possible that a longer exposure of the GO-coated grids to the oxidative air introduces more variation in how the Hfq-Crc-*amiE*$_{105}$ particles adhere to the graphene oxide monolayer.

A second notable observation relates to the local resolution estimates of the cryo-EM map (Figure 14B). From these estimates it is apparent that the cryo-EM density for the Crc molecules and parts of the RNA is resolved at

significantly lower resolutions than the rest of the assembly. This observation hints towards conformational heterogeneity of the assembly, which will be further studied in section 3.4.2.1.

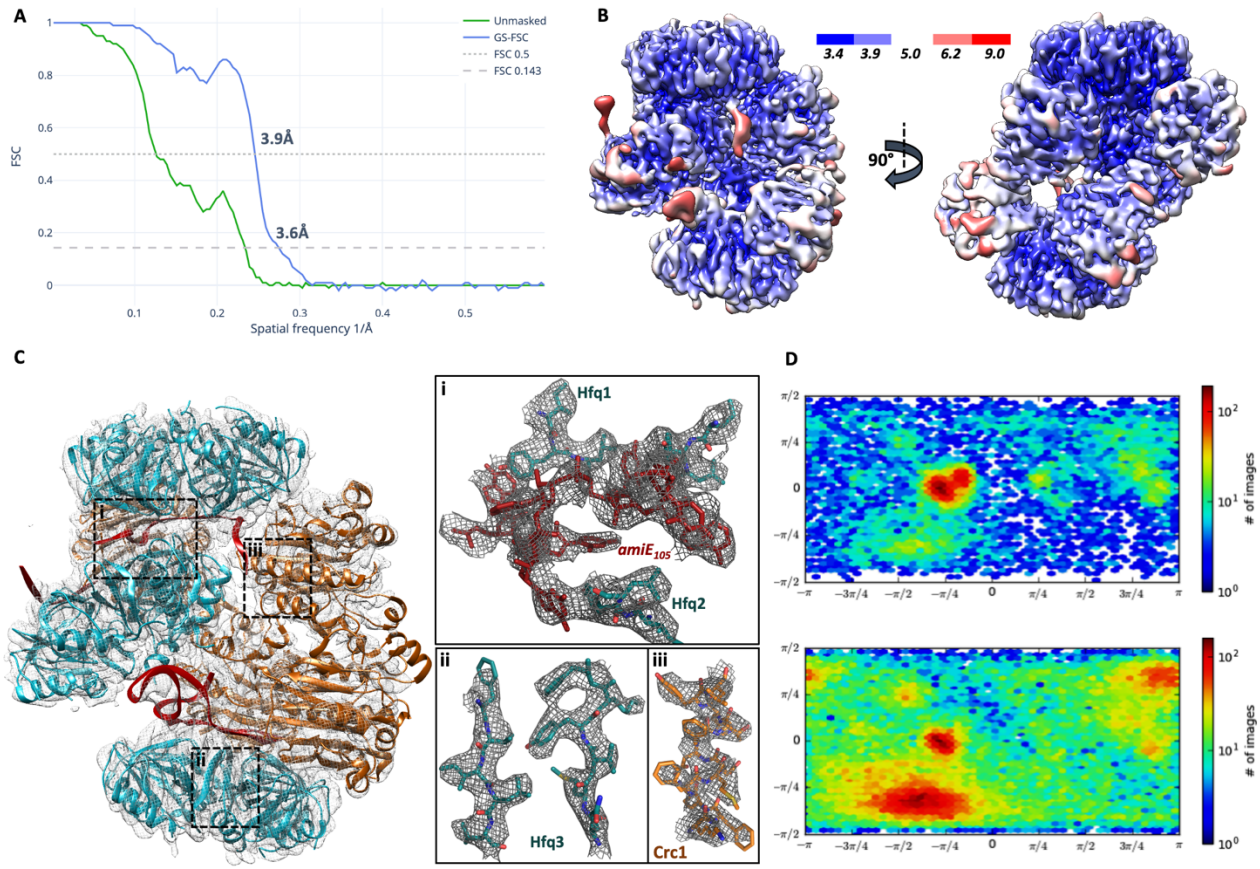


Figure 14: ***Resolution estimates of the cryo-EM reconstructions and model fitting. (A)*** *Global resolution estimates as calculated from independently reconstructed half maps via Fourier shell correlation (FSC).* ***(B)*** *Local resolution estimates were calculated in cryoSPARC ($FSC_{0.5}$). Some of the Crc molecules have significantly lower resolution estimates than the rest of the assembly, which could point to conformational heterogeneity.* ***(C)*** *The model fits well into the cryo-EM density as depicted here for various regions, including interaction interfaces between (i) $amiE_{105}$ and Hfq, (ii) Hfq beta strands and (iii) Crc alpha helices.* ***(D)*** *Projection distribution plots of the aligned particles from each of the two datasets collected on graphene oxide grids aged 24 hours (upper panel) and 48 hours (lower panel), revealing a discrepancy in orientation diversity in the thin, vitreous ice.*

3.4.2.1 3D variability analysis and focused refinements

To model potential conformational heterogeneity within the Hfq-Crc-$amiE_{105}$ assembly, the aligned particle set was subjected to 3D variability analysis (3DVA) in cryoSPARC (Punjani & Fleet, 2020). During 3DVA the resolution was limited to 6 Å, i.e. only structural variability up to 6 Å was considered, and six modes of variability were solved. Two modes of molecular motion were readily interpretable (Figure 15). The first mode describes the complex relative motions of two rigid body halves of the Hfq-Crc-$amiE_{105}$ assembly (Figure 15A and Figure 16, body 1 & 2). In particular, a first rigid body containing Hfq3-Crc2-Crc3-$amiE_{105}$ follows a screwdriver-like trajectory as opposed to the second rigid body, comprising of Hfq1-Hfq2-Crc1-Crc4-$amiE_{105}$ (Figure 15A). Interestingly, the $amiE_{105}$ hairpin loop that binds the Hfq2 proximal side (Figure 14C, hairpin1) appears to alternate interactions with the proximal side of two neighbouring Hfq protomers in the hexameric ring (Figure 15A inset i, and C). The second mode of variability captures the large molecular motion of a second putative $amiE_{105}$ hairpin loop on the Hfq2 distal side

(Figure 15B and C, hairpin2). The extensive conformational heterogeneity of this second hairpin loop rendered its cryo-EM density in the averaged consensus map uninterpretable (Figure 13C). The molecular motions captured in both modes of variability reveal the conformational landscape of the Hfq-Crc-*amiE*$_{105}$ complex (Figure 15C).

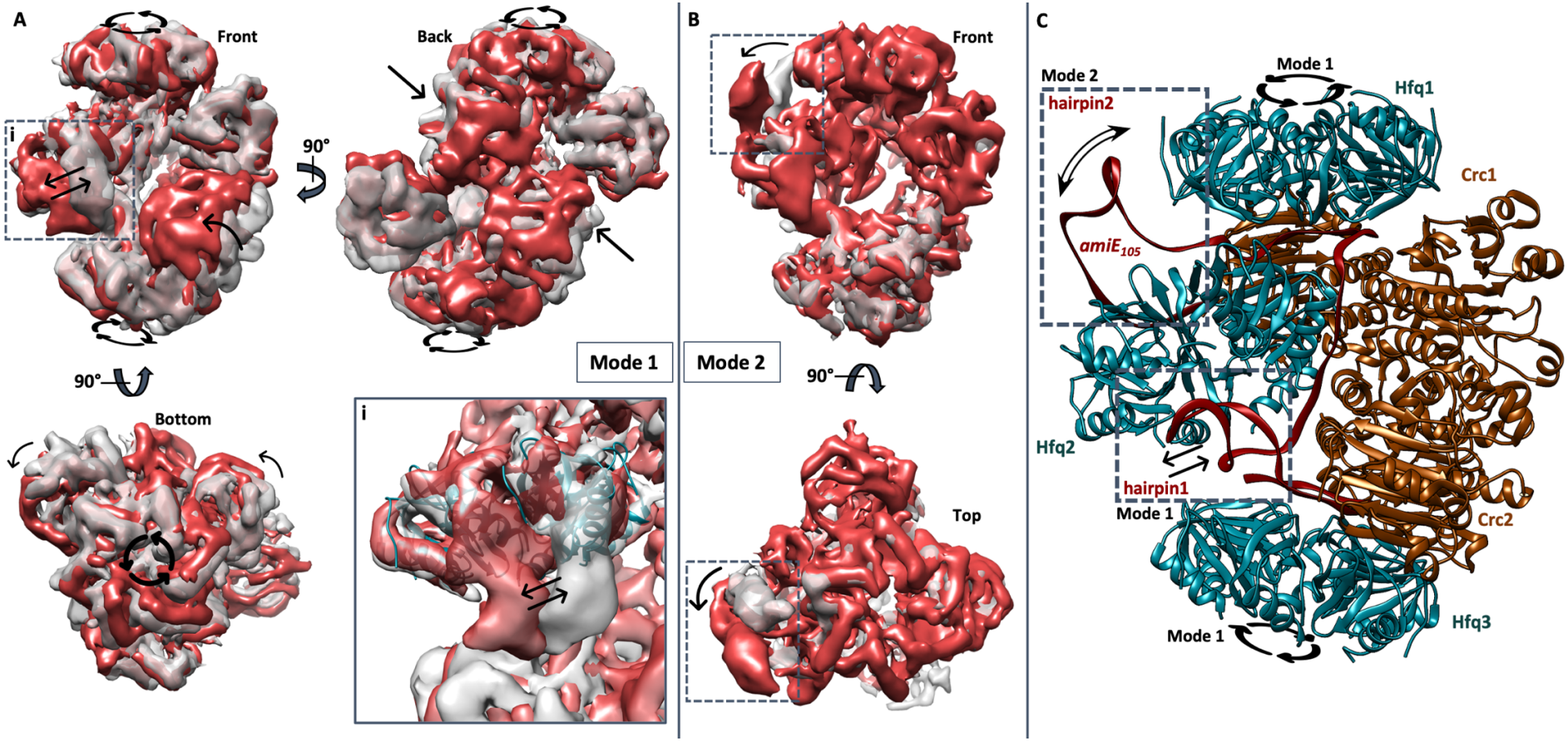


Figure 15: ***3D variability analysis of the Hfq-Crc-amiE$_{105}$ particles. (A)*** *The first mode of variability captures a complex 'screwdriver-like' motion within the Hfq-Crc-amiE$_{105}$ assembly, as highlighted by black arrows for the front, back and bottom views. Notably, the amiE$_{105}$ hairpin (hairpin1) appears to alternate interactions with two protomers of the Hfq2 hexamer (inset i). Molecular motions are depicted by overlayed grey and red maps, corresponding to the first and last frame of the encoded variability, respectively.* ***(B)*** *The second mode of variability captures the extensive 'rocking' motion of the second amiE$_{105}$ hairpin (hairpin2).* ***(C)*** *Conformational landscape of the Hfq-Crc-amiE$_{105}$ model as captured by both modes. All maps are lowpass filtered to 8 Å for interpretability.*

Based on the first mode of molecular motion within the Hfq-Crc-*amiE*$_{105}$ complex, the assembly was split in two rigid bodies for masked local refinements in cryoSPARC (Figure 16). This local refinement approach resulted in better resolved cryo-EM maps at 3.5 Å (body 1) and 3.4 Å (body 2) with significant improvements in the Crc-regions. Unfortunately, the map regions for the two *amiE*$_{105}$ hairpins did not improve, most likely due to significant residual conformational heterogeneity.

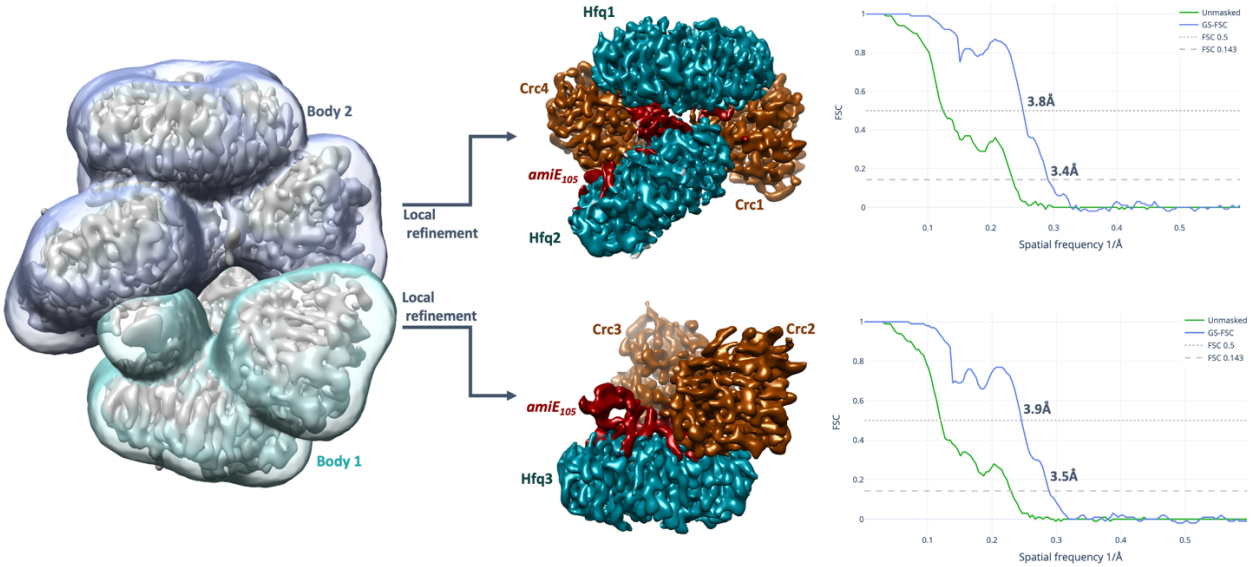


Figure 16: ***Rigid body definition and masked local refinements of the Hfq-Crc-amiE$_{105}$ assembly.*** *Soft masks were generated to define two rigid bodies based on 3D variability analyses of the Hfq-Crc-amiE$_{105}$ particles (Figure 15, mode 1). Subsequent masked local refinements of both rigid bodies in cryoSPARC resulted in better resolved Hfq and Crc densities, with global resolutions increasing to 3.5 Å and 3.4 Å for bodies 1 and 2, respectively (GS-FSC).*

3.4.3 Architectural principles of the Hfq-Crc-*amiE*$_{105}$ assembly

The Hfq-Crc-*amiE*$_{6ARN}$ cryo-EM structures that were solved in section 3.2 and published in Pei *et al.* (2019) provided many insights into how Hfq presents ARN-repeats to Crc but are limited by the short length of the *amiE*$_{6ARN}$ construct. In Pei *et al.*, a model for a Hfq-Crc assembly on the full *amiE* transcript was proposed based on the Hfq-Crc-*amiE*$_{6ARN}$ cryo-EM structures and a schematic comparison between the proposed and experimental Hfq-Crc assemblies is presented in Figure 17B. In the predicted model the full length *amiE* was thought to thread from one Hfq distal side through a basic half-channel formed by the central Crc dimer to fold into a second Brennan crown on the second Hfq distal side (Figure 17B, schematic on the right). The experimental Hfq-Crc-*amiE*$_{105}$ structure solved in this section shows that the predicted model was relatively accurate (Figure 17B, schematic on the left). *amiE*$_{105}$ is also sandwiched between two Hfq distal sides, and the target RNA threads from one distal side (Hfq3) to the other (Hfq1), forming Brennan crown-like folds on both. However, four Crc partners are stacked in the *amiE*$_{105}$ complex, and a third Hfq molecule (Hfq2) is wrapped in the centre of the assembly (Figure 17A and B, schematic on the left). In the next sections, architectural principles of the Hfq-Crc-*amiE*$_{105}$ translation-repression assembly will be discussed, including RNA recognition and presentation by Hfq (section 3.4.3.1), RNA-Hfq binding by Crc (section 3.4.3.2) and Crc dimerization (section 3.4.3.3).

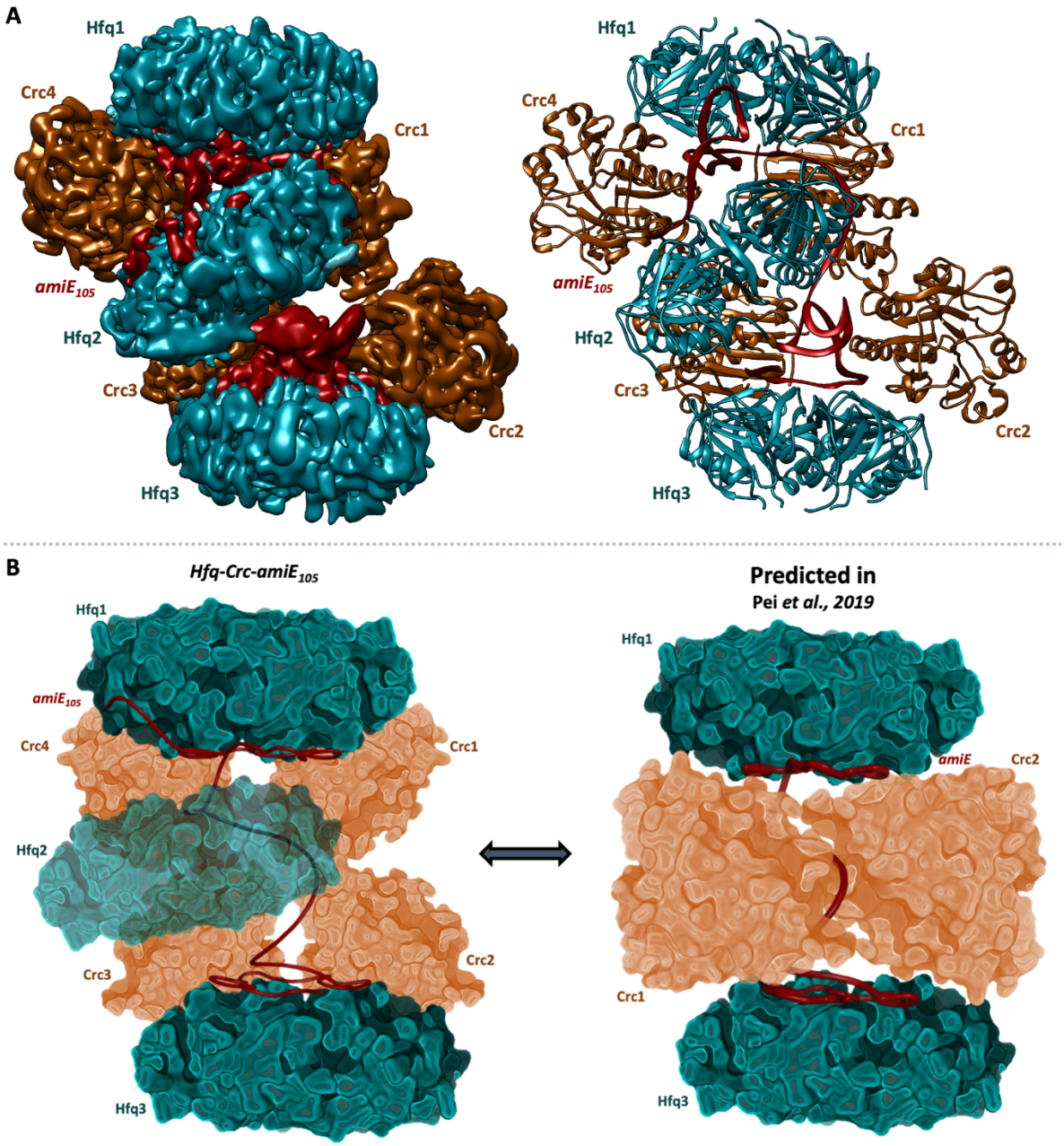


Figure 17: ***Schematic overview of Hfq-Crc-amiE$_{105}$ architecture. (A)*** *Colour coded map (left) and model (right) for the Hfq-Crc-amiE$_{105}$ translation-repression complex. Three Hfq chaperones present amiE$_{105}$ to four Crc molecules, forming a compact translation-repression assembly. amiE$_{105}$ interacts with all three Hfq-distal sides, and with the rim and proximal side of Hfq2.* ***(B)*** *Comparison between simplified schematics of the experimental Hfq-Crc-amiE$_{105}$ model (left) and the predicted model based on the cryo-EM studies in the first half of this chapter (right, Pei et al., 2019).*

3.4.3.1 amiE$_{105}$ recognition and presentation by Hfq

It is challenging to align the sequence of the *amiE$_{105}$* target in the model due to conformational heterogeneity of the Hfq-Crc-*amiE$_{105}$* assembly. However, the findings from section 3.2 combined with the sub-3.5 Å local resolution of the Hfq distal sides, the known ARN-rules and the predicted *amiE$_{105}$* secondary structure (mfold, Zuker, 2003) allowed for confident mapping of fragments of the *amiE* sequence. The *amiE$_{6ARN}$* motif at the 5'-end of *amiE$_{105}$* is bound by the Hfq3 distal side, folding into a first and complete Brennan crown (Figure 18A, Hfq3). Interestingly, in

the last ARN-repeat of this Brennan fold, the A-site is skipped, as shown by a lack cryo-EM density for an A-site base (Figure 18A and B, highlighted with *). The latter is the first example of a 'skipping'-violation of the ARN-rule of RNA binding on the Hfq distal side (Link *et al.*, 2009). Downstream of the *amiE$_{6ARN}$* region the RNA forms a short hairpin-loop which interacts with the proximal side of the central Hfq2 ring (Figure 18A and B, hairpin1), which will be further discussed in the next paragraph. The *amiE$_{105}$* then threads further over the Hfq2 rim to fold into a nearly complete Brennan crown fold on the Hfq1 distal side (Figure 18A, Hfq1). From the Hfq1 distal side *amiE$_{105}$* transitions continuously into a third Brennan crown fold on the Hfq2 distal side. Here *amiE$_{105}$* lifts off to form a second, mobile hairpin-loop structure (Figure 18A and B, hairpin2, Figure 15B), which folds back to occupy a final ARN-pocket on the same Hfq2 distal side. In the latter two Brennan folds, a series of 'mismatch'-violations of the ARN-rule are observed. Such 'mismatch'-violations are reasonably mild as they mainly manifest in pyrimidine bases occupying A- or R-sites on the Hfq distal side (Figure 18, bottom sequence, light blue letters). These observations indicate that in an integrated assembly like the Hfq-Crc-*amiE$_{105}$* complex more degenerate ARN-repeat motifs can be recognised and presented. Overall, the proposed *amiE$_{105}$* structure fits the predicted secondary structure reasonably well, with sensible annotations for ARN-motifs and hairpin-loop structures (Figure 18B).

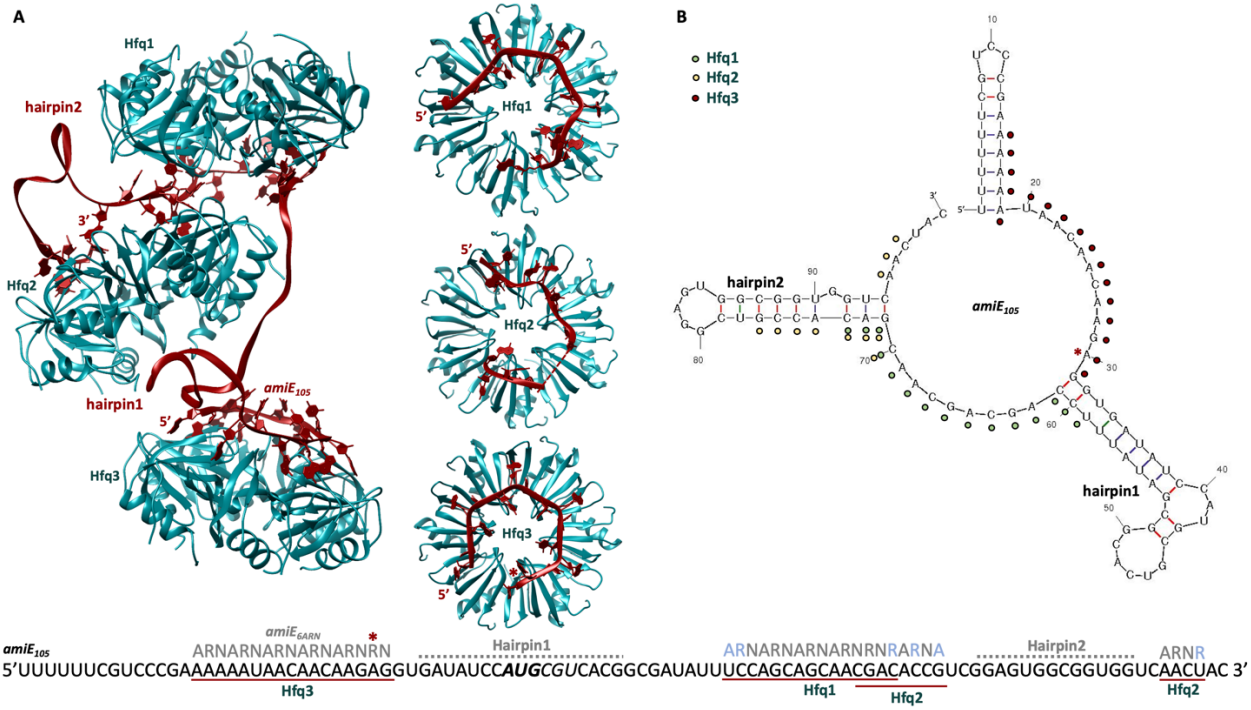


Figure 18: ***amiE$_{105}$ recognition and presentation by Hfq. (A)*** *amiE$_{105}$ is presented by three Hfq hexamers and adopts complete or partial Brennan crown folds on each distal side. The proximal side of Hfq2 coordinates an amiE$_{105}$ hairpin- loop structure (hairpin1). A second hairpin-loop forms at the 3' end of amiE$_{105}$ on the Hfq2 distal side (hairpin2).* ***(B)*** *Annotated secondary structure prediction of amiE$_{105}$ (mfold; Zuker, 2003). Coloured dots indicate which Hfq distal side presents the ARN-rich motif in the Hfq-Crc-amiE$_{105}$ model. An annotated sequence is depicted at the bottom of the Figure. Sequences that were modelled in the cryo-EM map are underlined in red and the Hfq distal sides they bind to are labelled in green. Occupied A-, R- or N- sites are annotated in grey above each modelled sequence. * refers to an A-site 'skipping'-violation and light blue letters refer to 'mismatch'-violations of the ARN-rule and are explained in the main text. The ranges for hairpin1 and hairpin2 are arbitrary due to low local resolution in the corresponding map regions.*

Notably, Hfq1 and Hfq2 share an *amiE*$_{105}$ stretch between their distal sides (Figure 18B and Figure 19A). In this transition region the *amiE*$_{105}$ sequence deviates from the ARN-repeat motif, as the nucleobases alternately occupy A- and R-pockets on the Hfq1 and Hfq2 distal sides (Figure 19A). In other words, N-site bases relative to one Hfq distal side occupy R- or A-pockets on the other Hfq distal side and vice versa. This observation points toward a novel function of N-site bases in ARN-rich motifs, where they could help drive formation of effective higher order assemblies by closely connecting and organising different Hfq chaperones on an RNA target.

The Hfq-Crc- *amiE*$_{105}$ structure also confirms a recognition mode of the Hfq-proximal side that was recently discovered by the Brennan group (Orans *et al.*, 2020). Here, Orans *et al.* solved a crystal structure (pdbID 5UK7) of a B-DNA duplex bound to the Hfq distal side (Figure 19C) and postulated that dsRNA stems could be recognised in a similar way in riboregulatory assemblies. In the Hfq-Crc translation-repression assembly presented here, a strikingly similar interaction is observed for an A-form hairpin-loop (hairpin 1) in the target RNA, confirming the hypothesis by Orans *et al.* (Figure 19B and C). The *amiE*$_{105}$ hairpin loop spans the Hfq hexamer and passes over the uridine-binding proximal-face pore. The local resolution in the cryo-EM map does not allow to accurately model Hfq sidechains involved in these interactions, except for Lys5 on the N-terminal tail of one Hfq protomer. The latter coordinates the phosphate backbone of the *amiE*$_{105}$ hairpin loop in a sequence-independent manner (Figure 19B). This is the first time dsRNA recognition and binding by the Hfq proximal side has been visualized in a biologically relevant assembly.

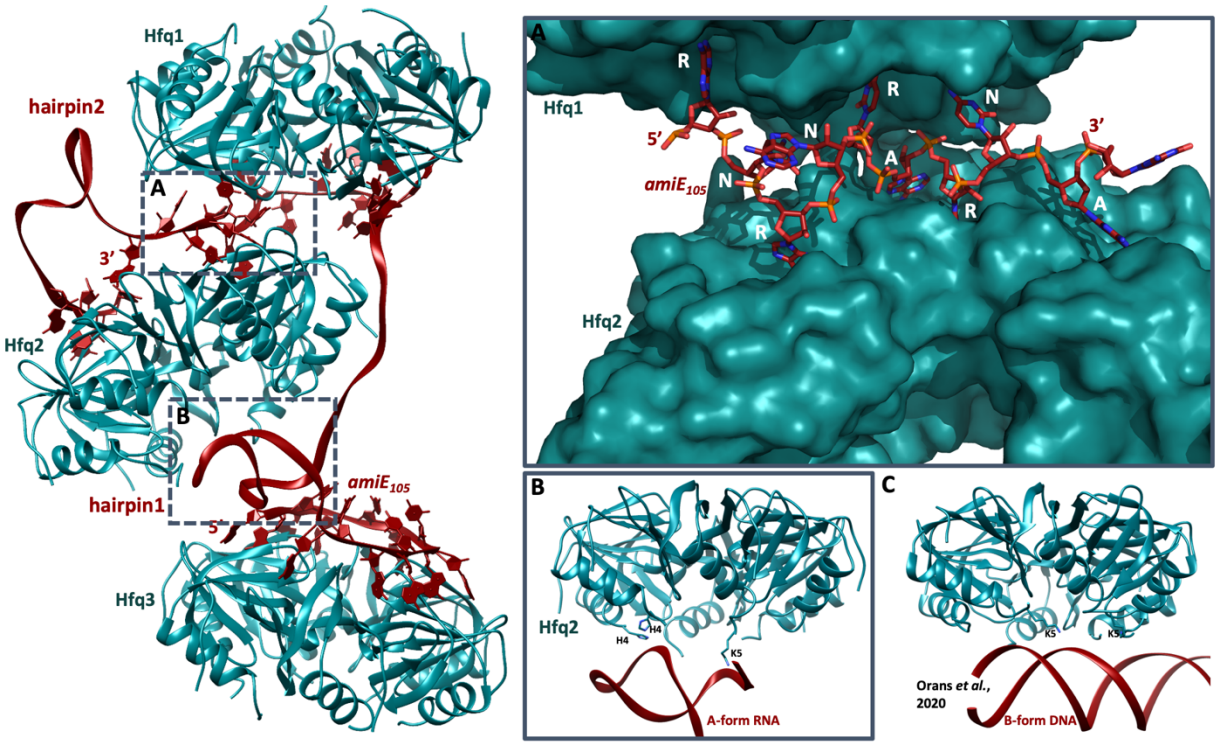


Figure 19: ***Brennan crown-sharing between Hfq distal sides and A-form RNA recognition by the Hfq proximal side. (A)*** *amiE*$_{105}$ *threads from the Hfq1 distal side to the Hfq2 distal, alternating A/R-sites between the two distal faces.* ***(B)*** *The Hfq proximal side binds an A-form RNA hairpin in the amiE*$_{105}$ *RNA molecule. The cryo-EM density in this region is weak due to conformational heterogeneity and no sidechain or base densities could be resolved, except for Lys5 on the N-terminal tail of a Hfq protomer. In addition, His4 is in close proximity to the amiE*$_{105}$ *A-form backbone.* ***(C)*** *Secondary structure recognition by the Hfq proximal side*

was recently proposed by Orans et al. (2020), who solved a crystal structure of Hfq in complex with B-DNA. Here too Lys5 appears to help coordinate the phosphate backbone.

3.4.3.2 Crc utilizes a preferred RNA binding mode in the amiE$_{105}$ complex

Hfq1 and Hfq3 each present *amiE$_{105}$* to two Crc molecules, forming identical Hfq-Crc-Crc core units (Figure 20A and B). Interestingly, each Crc partner molecule in the assembly interacts with the presented Brennan folds via the same interface. In particular, the Crc-*amiE$_{105}$* interaction mode is identical to that of Crc3 and Crc4 of the *amiE$_{6ARN}$* structures presented in section 3.2.3 (RNA binding mode 1). A detailed depiction of this Crc-binding can be found in Figure 9A. This preferred Crc-*amiE* interaction is in good agreement with the *in vivo* reporter assays discussed in section 3.2.4 (Figure 8B, R140E mutant and Figure 9A). An additional polar contact is made between a His131 of Crc1, located on a signature alpha helix, and U57 of *amiE$_{105}$* (Figure 20C).

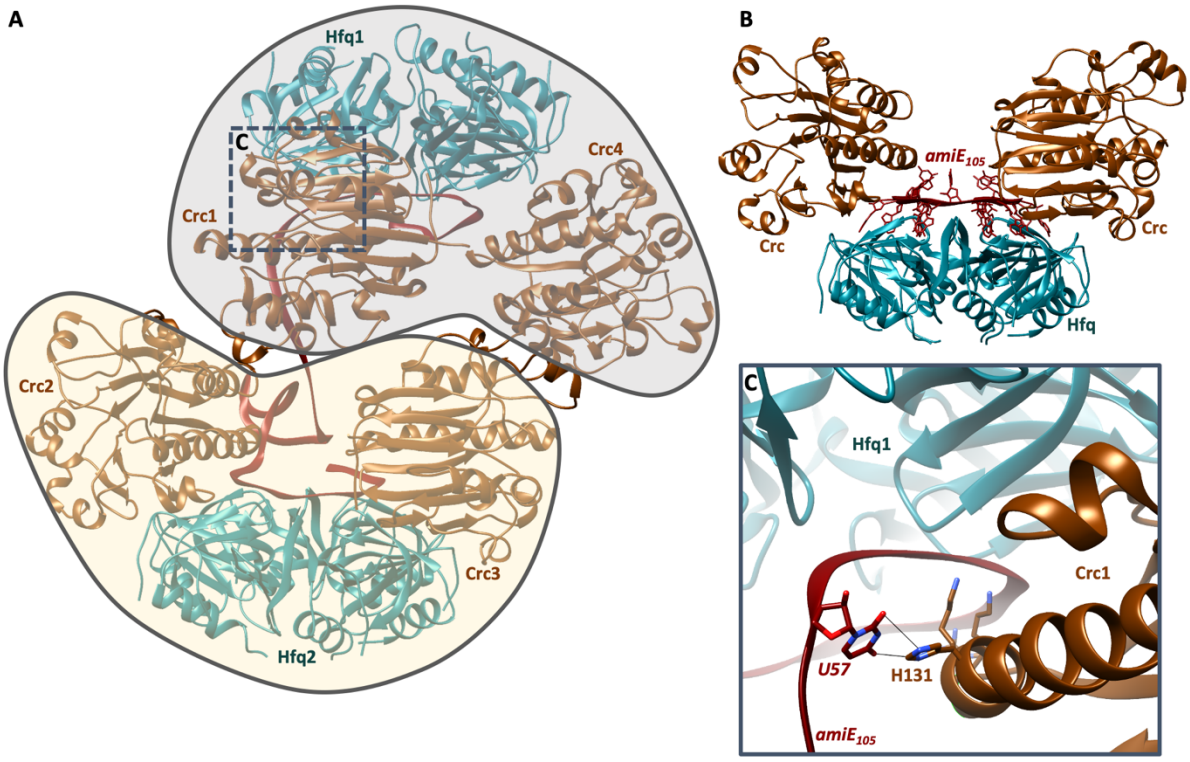


Figure 20: ***Crc utilises a single binding mode to interact with the presented Brennan crown folds. (A)*** *At the basis of the Hfq-Crc-amiE$_{105}$ assembly are two Hfq-Crc-Crc core units that interconnect to form a higher order assembly.* ***(B)*** *Model of the Hfq-Crc-Crc core unit. The Crc binding mode is identical to that of Crc3 and Crc4 in Figure 9A.* ***(C)*** *An additional polar contact is formed between Crc1 (H131) and amiE$_{105}$ (U57). Density for U57 does not allow for accurate nucleobase modelling and the depicted orientation is putative.*

3.4.3.3 Permissive Crc dimerization interfaces help drive higher order assembly

Crc is monomeric in solution (Milojevic *et al.*, 2013), yet in the assemblies that have been described so far Crc appears to self-interact and form homodimers. In the Hfq-Crc-*amiE$_{105}$* translation-repression complex Crc forms two distinct dimers upon engaging the presented RNA substrate (Figure 21A). These dimers not only differ in their dimerization interface but also have significantly different relative orientations of the Crc molecules (Figure 21B, dimers 1 and 2). Strikingly, both Crc dimers also differ significantly from the partially artificial Crc dimers in the Hfq-

Crc-*amiE*$_{6ARN}$ assemblies, which are described in section 3.2.4 (Figure 21B, 2:2:2 and 2:4:2 dimers). Dimer interface 2 in Figure 22 is in good agreement with the *in vivo* reporter assays described in section 3.2.5 (Figure 9B, E193R mutant), which further confirms the *in vivo* relevance of the Hfq-Crc-*amiE*$_{105}$ model. These observations indicate that Crc dimerization is promiscuous and can be constituted over different permissive dimerization interfaces, but only under conditions where Crc is preorganised.

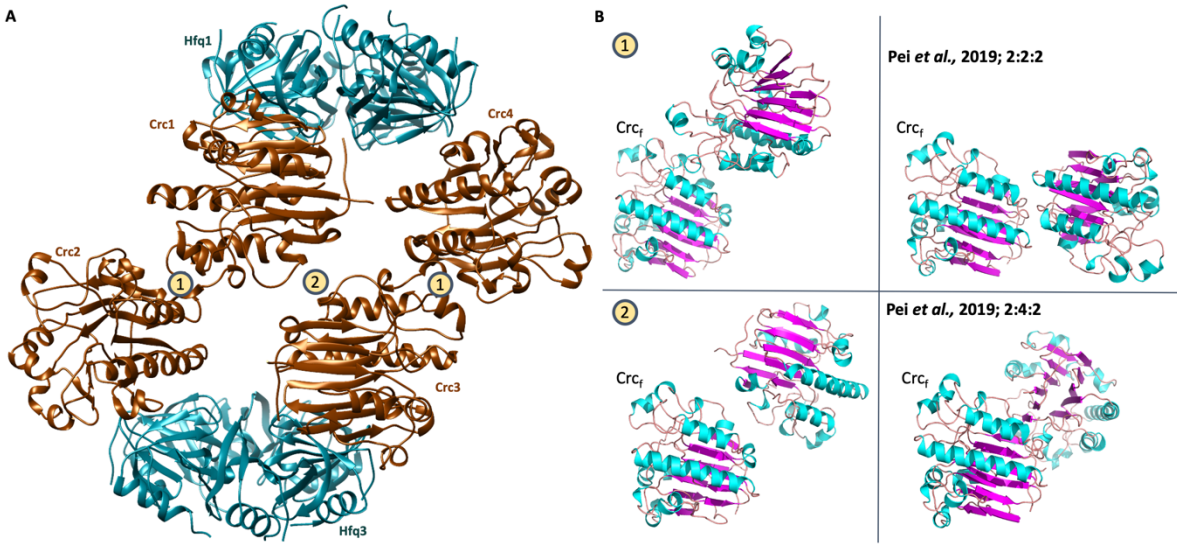


Figure 21: ***Crc dimers in the Hfq-Crc-amiE$_{105}$ assembly. (A)*** *Two different Crc dimer interfaces are observed in the Hfq-Crc-amiE$_{105}$ assembly (highlighted with yellow circles 1 and 2).* ***(B)*** *Comparison between Crc dimers 1, 2 and the two dimers described in Figure 9C (section 3.2 and Pei et al., 2019). For comparison, each dimer is depicted relative to a reference Crc, Crc$_f$, which is fixed in space for all four dimers. It is apparent that Crc can form several homodimers via different permissive dimerization interfaces to drive higher order assembly formation.*

3.4.4 Intermediate complexes hint towards a defined and cooperative assembly pathway

It is clear that the first step in assembly of the *ami*E translation-repression complex is the recognition and presentation of the RNA target by Hfq. As mentioned in section 3.4.2, many different assemblies were observed during grid screening. Although unexpected, these intermediate assemblies could help answer a key question: do all three Hfq molecules have to cooperatively present the mRNA target to Crc for the translation-inhibition assembly to form? In this scenario Hfq must 'saturate' the nascent RNA during transcription, after which Crc can bind the Hfq-RNA intermediate to prevent translation, much like a two-step process. Alternatively, Crc could bind ARN-rich regions of the nascent transcript as they are presented 'on the fly' by Hfq (Figure 30). To answer this question cryo-EM samples were prepared at three different molar ratios of Hfq and Crc. The samples had a fixed molar ratio for Hfq and *amiE*$_{105}$, i.e. two Hfq molecules per *amiE*, and varied only in the concentration of Crc molecules. Although three Hfq molecules are expected to bind *amiE*$_{105}$, working at sub-stoichiometric Hfq concentrations allowed for lower order intermediates to be studied (Figure 22). For each of the three samples a dataset was collected on a 300 kV Titan Krios (FEI) equipped with a K3 detector (Gatan) (BioCem, Department of Biochemistry, University of Cambridge). Motion correction of the movie frames, CTF estimation/correction and particle picking were performed in Warp (Tegunov & Cramer, 2019) and 2D/3D alignments and averaging were carried out in cryoSPARC (Punjani *et al.*, 2017).

Careful 2D classification analyses of the three different cryo-EM datasets revealed a wide range of lower order intermediate assemblies. In Figure 22, these intermediates are sorted in a visual 2D table according to the input sample (rows) and the observed complexity (columns). Due to preferred orientation of the particles in the vitreous ice and conformational heterogeneity, 3D reconstructions were not always possible. However, prior structural knowledge of the sample components allowed for accurate annotation of the 2D class averages. The lowest order intermediate observed on the grids consists of a single Hfq hexamer and occurred in all three samples (Figure 22, tiles 1, 6 & 11). At low molar ratios of Crc in the sample (one or four Crc molecules per $amiE_{105}$), two other Crc-less assemblies are observed, where two (Figure 22, tiles 2 & 7) or three (Figure 22, tiles 3 & 8) Hfq molecules can assemble on the $amiE_{105}$ target. All three samples also contained an assembly intermediate where a single Crc molecule and two Hfq chaperones assemble on $amiE_{105}$ (Figure 22, tiles 4, 9 & 12).

When a 11-fold excess of Crc was added to the sample (Figure 22, bottom row), the most abundant class contained two Crc molecules bound to a single Hfq-$amiE_{105}$ intermediate (Figure 22, tile 13). 3D reconstructions of this intermediate resulted in a 3.2Å cryo-EM map (GS-FSC), revealing a Hfq-Crc-Crc core unit of the translation-repression assembly, as described in section 3.4.3.2 (Figure 23, left inset and Figure 20B). In addition, the RNA density at the Hfq distal side corresponds to the 5' $amiE_{6ARN}$ motif. It is therefore likely that this lower order intermediate serves as an initial scaffold for further co-transcriptional assembly of a full translation-repression complex in the cell. Indeed, a different subset of particles corresponds to what is likely to be the next intermediate in the assembly process, as revealed by a 3.9 Å reconstruction (Figure 23, middle inset). In this sub-assembly two Hfq chaperones present $amiE_{105}$ to three Crc molecules. Lastly, 2% of the particles correspond to the final assembly, as described in section 3.4.3, closing the putative Hfq-Crc-$amiE_{105}$ assembly pathway (Figure 23, right inset).

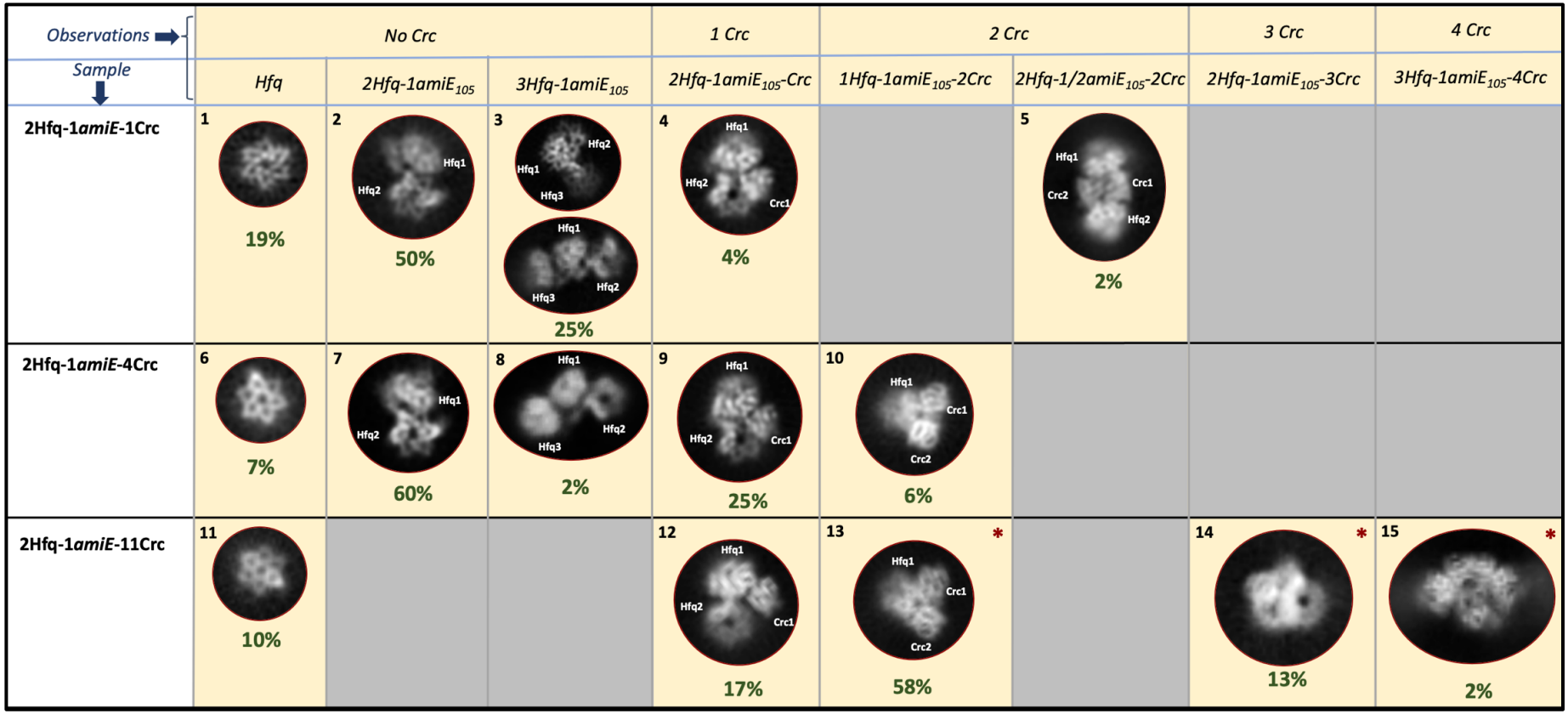


Figure 22: ***Gallery of Hfq-Crc intermediates.*** *The first column contains the molar ratios of the protein and RNA components in the cryo-EM sample. The top two rows describe the observations after extensive 2D/3D classification rounds. The boxes with a * correspond to assemblies for which a high-resolution 3D reconstruction was generated. For each class, the number of contributing particles is depicted as a fraction of the clean particle stack.*

Taken together, these observations reveal a probable cooperative assembly pathway of the Hfq-Crc complex as it would form on a nascent *amiE* transcript in the cell. It should be noted that the intermediates depicted in Figure 23 only form at higher concentrations of Crc. Interestingly, a scenario where three Hfq molecules together present the whole $amiE_{105}$ transcript, prior to Crc binding, is possible as well. This is evidenced by the formation of such intermediates at low molar ratios of Crc (Figure 22, tiles 3 and 8). Therefore, alternative assembly pathways cannot be ruled out and are not necessarily mutually exclusive in the cell. It is plausible that the overall Hfq-Crc-*amiE* assembly process corresponds to an assembly funnel with multiple possible trajectories to the full translation-repression assembly.

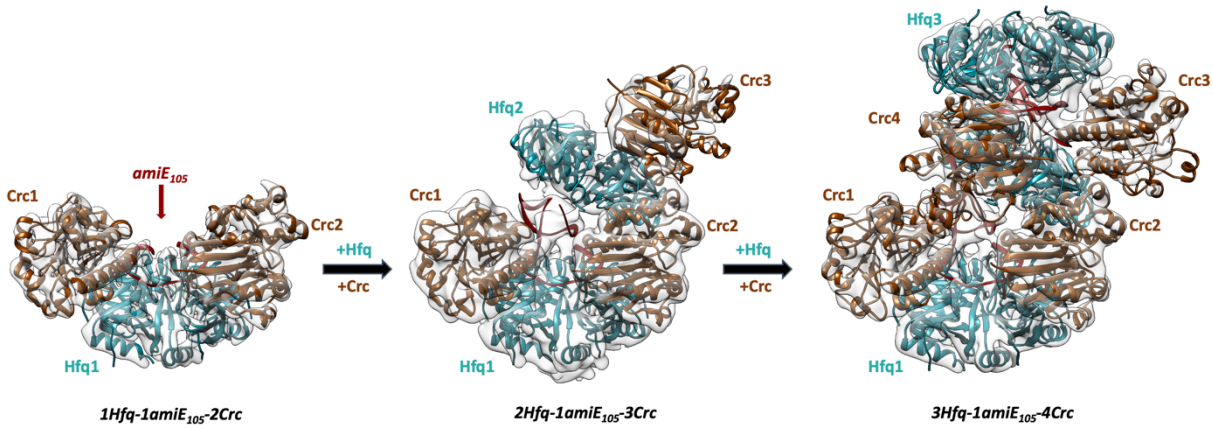


Figure 23: ***A cooperative assembly pathway for Hfq-Crc assemblies on amiE.*** *The $amiE_{6ARN}$ region is bound by the Hfq distal side and presented as a Brennan crown. Subsequently, two Crc molecules recognise and engage the Hfq-RNA complex (left). An additional Hfq and Crc can then bind to form a higher order intermediate (middle). In a final step, a third Hfq engages the intermediate assembly, together with a fourth Crc molecule (right), to fully mask the amiE 5'-end from ribosomes in the cell. All maps are lowpass-filtered to 7 Å for interpretability*.

3.4.5 Cryo-EM analysis of the Hfq-Crc-rbsB$_{110}$ *assembly*

There are three architectural principles that can be postulated from the Hfq-Crc-*amiE*$_{105}$ model described in sections 3.4.2 and 3.4.3. Firstly, sequence specific RNA target binding is ensured by the Hfq distal side, which prefers ARN-rich motifs. Secondly, the Hfq proximal side can recognise and interact with secondary structure elements in the target RNA fold, such as hairpin loops. Both these features are likely to determine the number of Hfq chaperones involved in a given translation-repression assembly and its quaternary structure. Thirdly, Crc can drive assembly of translation-repression complexes on a diverse pool of mRNA targets via its many permissive dimerization interfaces. To test the universality of the principles, cryo-EM studies were carried out on the Hfq-Crc-*rbsB*$_{110}$ assembly. As discussed in section 3.4.1, *rbsB* encodes a putative ribose transporter and was found to be one of the more than 100 mRNA targets that are co-regulated by Hfq and Crc (Kambara *et al.*, 2018).

Cryo-EM samples were prepared at an optimised molar ratio of 7Hfq-1*rbsB*$_{110}$-22Crc and applied to Quantifoil grids coated with a graphene oxide monolayer (Materials and Methods sections 5.4 and 5.5). Preliminary datasets indicated that the particles suffered from preferred orientations in the vitreous ice. Therefore, a dataset was collected at 35° stage-tilt to avoid anisotropic 3D reconstructions (Tan *et al.*, 2017) (Titan Krios, FEI, K3 detector, Gatan, BioCem, Department of Biochemistry, University of Cambridge). Motion correction of the movie frames, CTF estimation/correction and particle picking were carried out in Warp (Tegunov & Cramer, 2019), 2D/3D alignments and averaging were carried out in cryoSPARC (Punjani *et al.*, 2017, Punjani *et al.*, 2019). A representative cryo-EM image at 2.5 µm defocus is depicted in Figure 24A. Several rounds of 2D classification resulted in clean 2D class averages with interpretable secondary structure elements (Figure 24B). 3D reconstructions from a clean particle set reveal a compact assembly where rbsB is presented by two Hfq hexamers to three Crc molecules (Figure 24C and E). Interestingly, diffuse additional density for a third Hfq hexamer engaging the assembly with its proximal side can be observed at low thresholds (Figure 24D, Hfq3). Conformational heterogeneity is the root cause for this poorly resolved Hfq3 density and will be discussed in section 3.4.5.1.

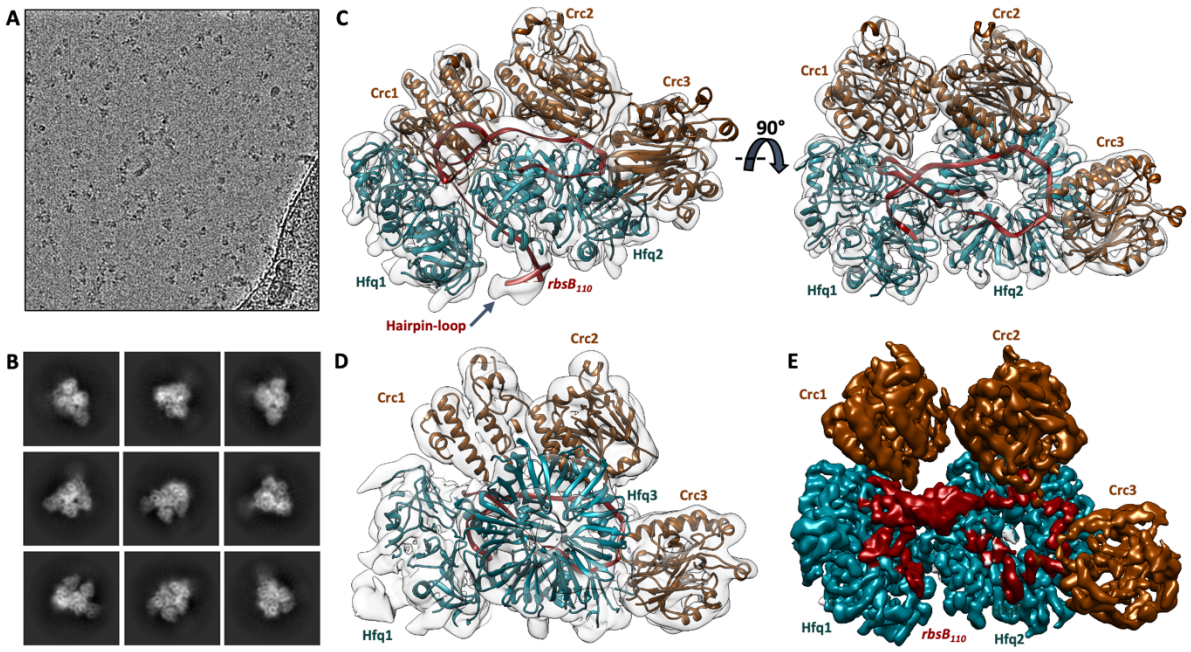


Figure 24: ***Cryo-EM reconstruction of the Hfq-Crc-rbsB$_{110}$ assembly. (A)*** *Representative image at 2.5 µm defocus, lowpass filtered to 20 Å.* ***(B)*** *Selection of 2D class averages.* ***(C)*** *At first sight, two Hfq hexamers and three Crc molecules assemble onto rbsB$_{110}$ to form a translation-repression complex.* ***(D)*** *At low density thresholds a third Hfq hexamer can be docked in the density (Hfq3), engaging the presented rbsB$_{110}$ RNA with its proximal side. The map and model in (D) are clipped along the z-axis for interpretability. The maps in (C) and (D) were filtered with LAFTER for interpretability (Ramlaul et al., 2020).* ***(D)*** *A 3D denoised, colour coded map of the Hfq-Crc-rbsB$_{110}$ assembly. 3D denoising was done with a Noise2Noise based algorithm as part of the Warp package (Lehtinen et al., 2018; Tegunov & Cramer, 2019).*

After 3D refinement and global/local CTF refinements a 3.8 Å (GS-FSC) consensus map was reconstructed (Figure 25A) with local resolution estimates ranging from 3.2 Å in the central map regions to 7.7 Å at the periphery (Figure 25C). The aligned particle orientations as plotted by cryoSPARC have the expected ring-like distribution that is typical for data collections with a tilted stage. At these resolutions Hfq and Crc structures could be readily docked and refined and large stretches of RNA could be traced and built *de novo* (Figure 25D). A comprehensive description of the Hfq-Crc-*rbsB$_{110}$* model will be given in section 3.4.6.

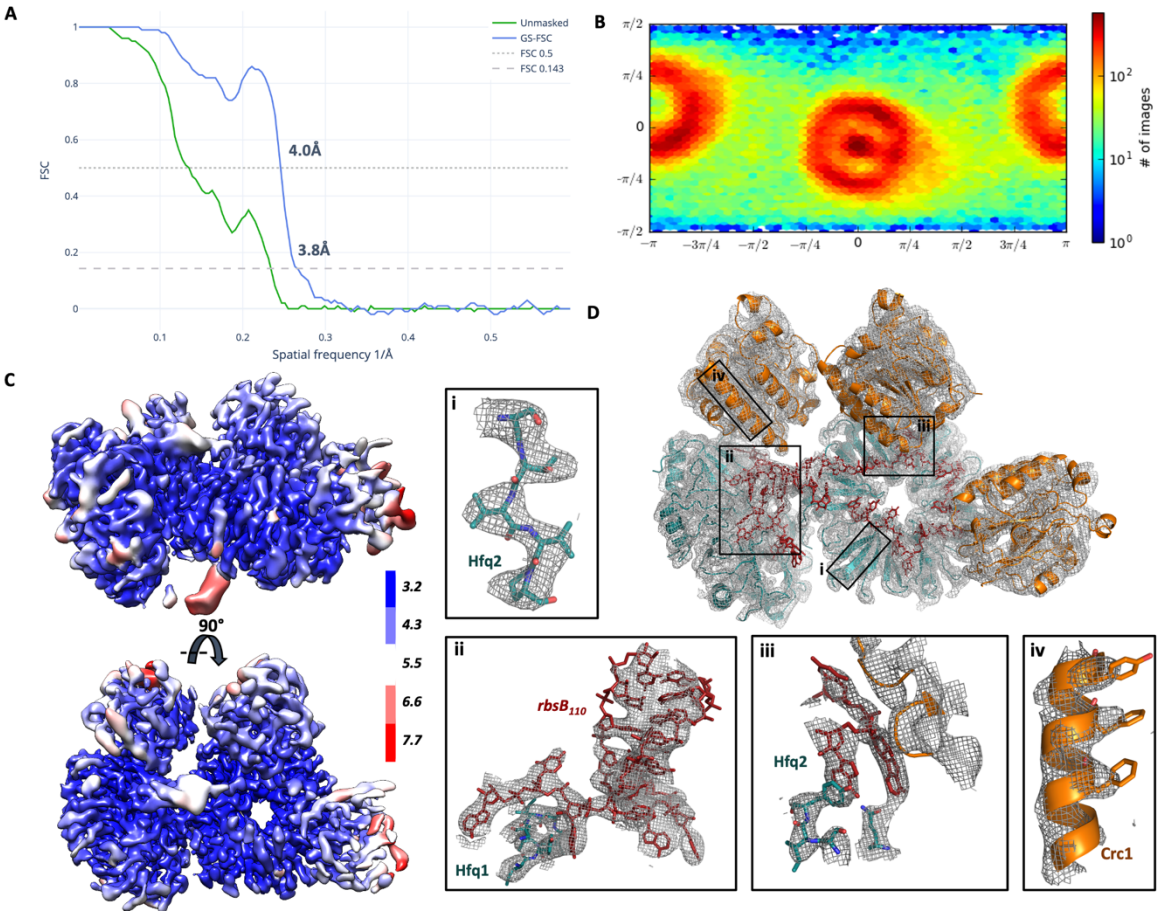


Figure 25: ***Resolution estimates of the Hfq-Crc-rbsB$_{110}$ cryo-EM reconstructions and model fitting.*** ***(A)*** *Global resolution estimates as calculated from independently reconstructed half maps via Fourier shell correlation (FSC).* ***(B)*** *2D presentation of the particle orientation distribution as generated in cryoSPARC.* ***(C)*** *Local resolution estimates were calculated in cryoSPARC (FSC$_{0.5}$) and range from 3.2 Å to 7.7 Å.* ***(D)*** *The model fits well into the cryo-EM density as depicted here for various regions, including Hfq beta strands (i), interaction interfaces between rbsB$_{110}$ and Hfq (ii and iii), and Crc alpha helices (iv). The map used for panel D was a 3D-denoised version of the refined map at 3.8 Å (Lehtinen et al., 2018; Tegunov & Cramer, 2019).*

3.4.5.1 3D variability analysis of the Hfq-Crc-rbsB$_{110}$ particles

Although local resolution estimates show that the majority of the refined map is resolved at sub-3.5 Å resolution, the global resolution estimate is considerably lower, at 3.8 Å. In particular, the cryo-EM density for Hfq3 is very weak (not shown in Figure 25). To assess whether these observations are due to conformational heterogeneity within the assembly, 3D variability analyses (3DVA) were carried out on the aligned Hfq-Crc-*rbsB$_{110}$* particles in cryoSPARC (Punjani & Fleet, 2020). Only structural variability up to 6 Å was considered during 3DVA, and six modes of molecular motion were solved. Two modes of molecular motion were readily interpretable (Figure 26). The first mode of variability captured subtle 'rocking'-like molecular motions by Crc1 and Crc2 (Figure 26A). In addition, a large 'rolling' motion of Hfq3 was encoded in the first mode of variability as well (Figure 26B), which is likely to be the cause for the corresponding diffuse cryo-EM density in the refined maps. Lastly, the second mode of variability reveals an extensive rocking motion of a hairpin-loop bound to the Hfq2 proximal side (Figure 24C and Figure 26C and D). As for the Hfq-Crc-*amiE$_{105}$* assembly, 3DVA helps understand the conformational landscape of the Hfq-Crc-

rbsB$_{110}$ complex and explains why some components of the assembly are poorly resolved in the averaged reconstructions.

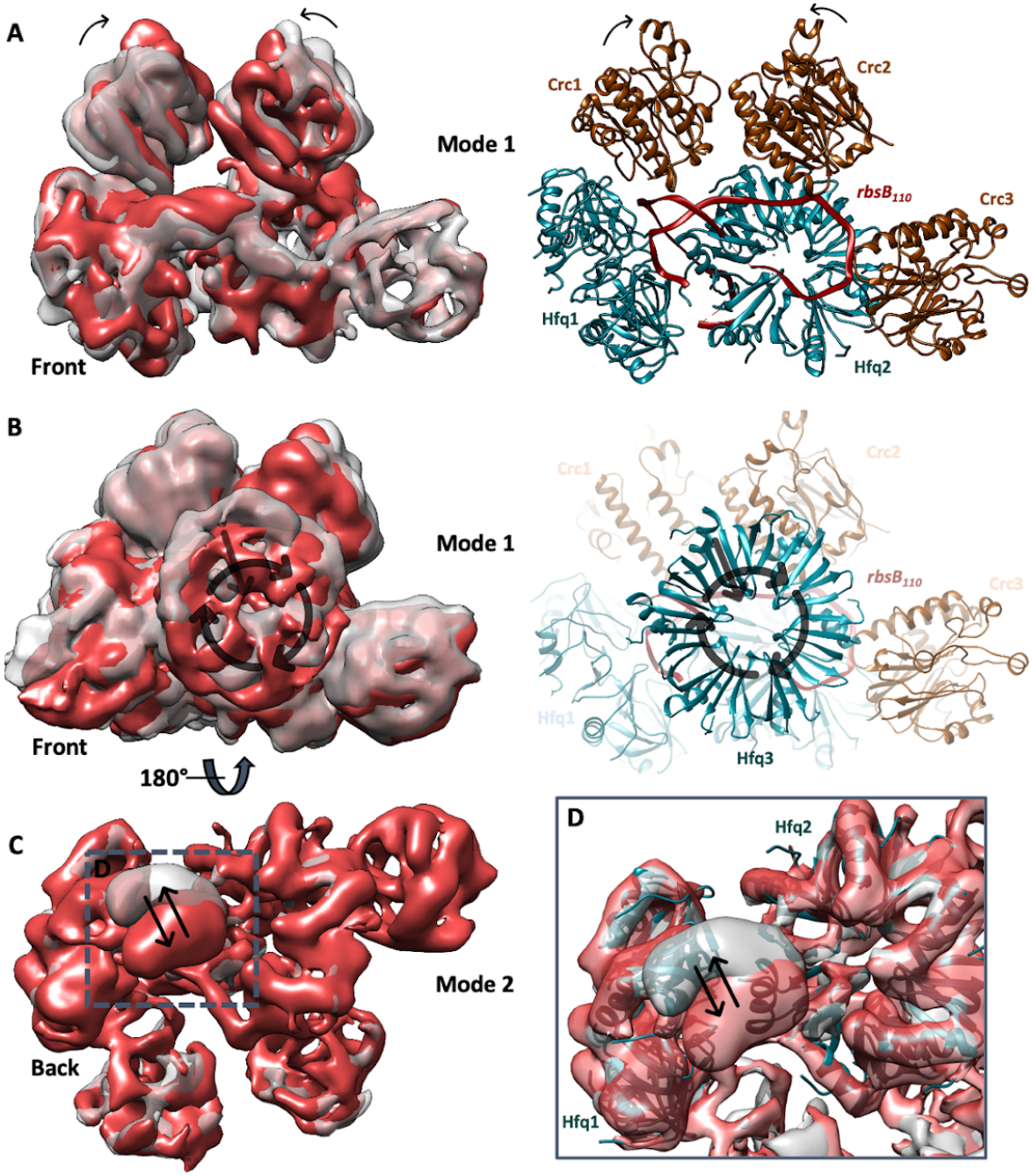


Figure 26: ***3D variability analysis of the Hfq-Crc-rbsB*$_{110}$ *particles.(A)*** *The first mode of variability captures a subtle rocking motion in- and out- of plane by the Crc1 and Crc2 molecules in the assembly, as highlighted by black arrows. Molecular motions are depicted by overlayed grey and red maps on the left, corresponding to the first and last frames of the encoded variability, respectively. On the right, the motions are annotated on the model.* ***(B)*** *The first mode of variability also captures the large rolling motion of Hfq3, explaining why the corresponding density in the refined consensus maps is diffuse.* ***(C)*** *The second mode of variability captures the extensive 'rocking' motion of a rbsB hairpin coordinated by the Hfq2 proximal side. All maps are lowpass filtered to 8 Å to aid visual interpretability.*

3.4.6 Architectural principles of the Hfq-Crc-rbsB$_{110}$ assembly

From the results described in section 3.4.5 it is clear that the quaternary structure of the *rbsB$_{110}$* translation-repression assembly is significantly different from the *amiE$_{105}$* translation-repression assembly. In this section a detailed overview of the Hfq-Crc-*rbsB$_{110}$* translation-repression complex is presented, including *rbsB* recognition and presentation by Hfq (section 3.4.6.1) and Crc dimerization (section 3.4.6.2).

3.4.6.1 rbsB$_{110}$ recognition and presentation by Hfq

As shown in section 3.4.5.1 the Hfq-Crc-*rbsB$_{110}$* translation-repression assembly is conformationally heterogeneous, resulting in poorly resolved regions of the map. In addition, the backbone for only 60 nucleotides out of 110 could be traced in the cryo-EM density. The cryo-EM density for the *rbsB* segments bound to the Hfq1 and Hfq2 distal sides, however, was resolved at sub-3.5 Å resolution. Together with the ARN-rule for RNA binding to the Hfq distal side and a secondary structure prediction of *rbsB$_{110}$*, putative sequence fragments could be aligned in the cryo-EM model (Figure 27). A nearly complete ARN-triplet repeat motif was modelled into the cryo-EM density on the Hfq2 distal side. Only one 'mismatch'-violation of the ARN-rule (Link *et al.*, 2009) was observed where the A-pocket is occupied by a pyrimidine nucleobase rather than adenine (Figure 27, Hfq2-binding segment in bottom sequence). The Hfq1 distal side, on the other hand, coordinates two short A-rich *rbsB* fragments, which were mapped upstream and downstream of the longer ARN-motif (Figure 27). These two annotations are a good match between the cryo-EM map and the *rbsB* sequence, with only one observed 'mismatch'-violation of the ARN-rule where a pyrimidine base occupies an R-pocket (Figure 27, bottom sequence, Hfq1a and Hfq1b segments). In addition, the cryo-EM density shows that an A-pocket is skipped on the Hfq1 distal side (Figure 27, highlighted with *), in good agreement with the mapped *rbsB* fragment. Mapping short ARN-motifs, however, is not unambiguous when flanking regions are not very well resolved, as is the case for the Hfq-Crc-*rbsB$_{110}$* cryo-EM map.

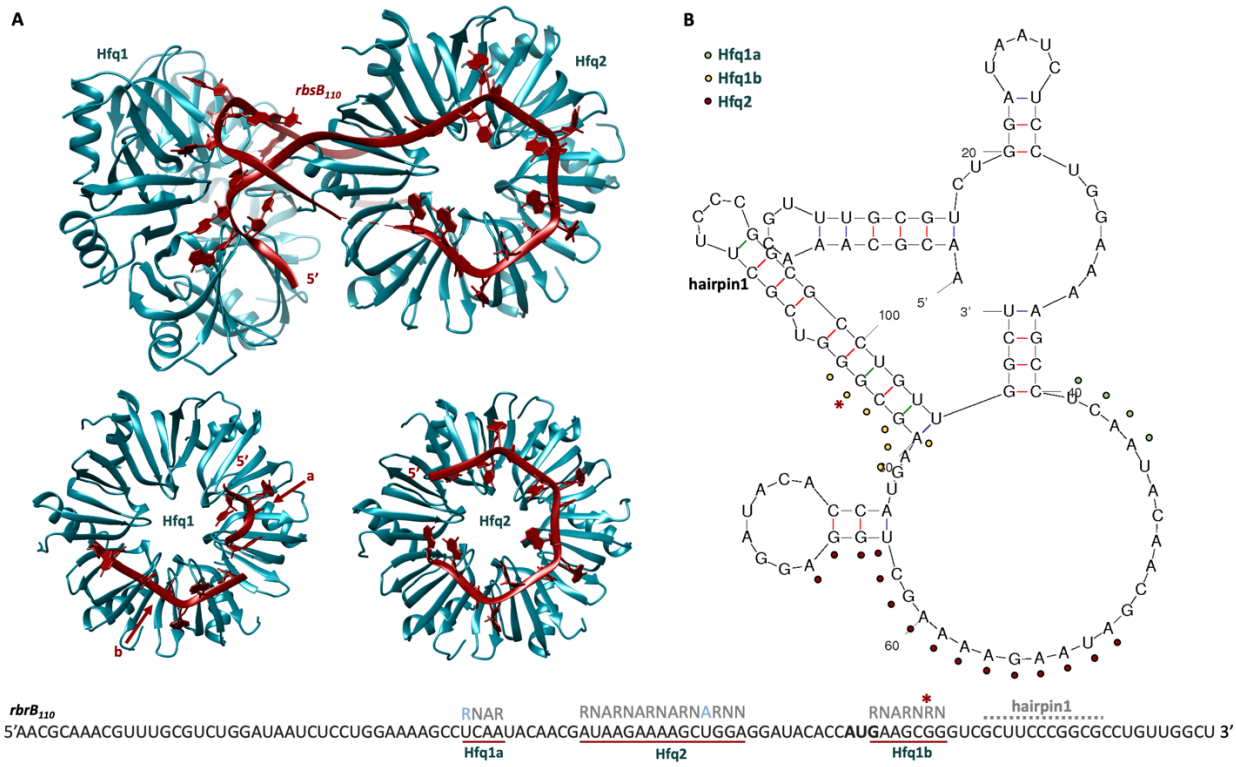


Figure 27: ***rbsB$_{110}$ recognition and presentation by Hfq. (A)*** *rbsB$_{110}$ is presented by two Hfq hexamers and adopts complete or partial Brennan crown folds on each distal side. The distal side of Hfq2 binds a nearly perfect and complete ARN-motif in the rbsB sequence. The Hfq1 distal side coordinates two putative, distant short A-rich motifs in the rbsB sequence (annotated with a and b)* ***(B)*** *Annotated secondary structure prediction of rbsB$_{110}$ (mfold; Zuker, 2003). Coloured dots indicate which Hfq distal side presents the ARN-rich motif in the Hfq-Crc-rbsB$_{110}$ model. An annotated sequence is depicted at the bottom of the figure. Sequences that were modelled in the cryo-EM map are underlined in red and the Hfq distal sides they bind to are annotated in green. Occupied A-, R- or N- sites are annotated in grey above each modelled sequence. * refers to an A-site 'skipping'-violation and light blue letters refer to 'mismatch'-violations of the ARN-rule and are explained in the main text. The proposed hairpin1 range at the 3' end is arbitrary due to the limited local resolution in the corresponding map region. Hfq3 is not depicted in panel A for interpretability.*

Interestingly, two secondary structure elements could be observed in the *rbsB$_{110}$* cryo-EM density. At the 3'-end of the modelled *rbsB$_{110}$* backbone, a short fragment of A-form RNA could be modelled in the density, likely corresponding to the stem of a small hairpin-loop (Figure 28A, Figure 27B, annotated as hairpin1). Due to significant molecular motion of this short hairpin loop the local resolution of the corresponding reconstruction is low (Figure 26C). As for the Hfq-Crc-*amiE$_{105}$* model, this hairpin-loop is likely to be coordinated by the Hfq N-termini on the Hfq2 proximal side, although these interactions could not be modelled. Strikingly, the cryo-EM density that would correspond to the loop of the hairpin structure is in close proximity to the N- and C-termini of a neighbouring Hfq1 protomer. Although not resolved in the cryo-EM map, it is likely that basic residues such as Arg19 and Arg66 on these termini help coordinate the hairpin-loop (Figure 28A).

A second structured element in the cryo-EM reconstruction is a sheet-like RNA duplex at the very centre of the assembly (Figure 28B). This RNA duplex has a high helical rise, i.e. a weak helical turn, and therefore does not

adopt an A-form fold. The quaternary structure of the Hfq-Crc-*rbsB*$_{110}$ complex suggests that the Hfq2 rim and Hfq3 proximal side could help coordinate this RNA duplex. However, no candidate basic surface patches were observed in close enough proximity to the RNA duplex, either for Hfq2 or Hfq3 in the assembly (Figure 28B). The rolling motion observed for Hfq3 in the *rbsB*$_{110}$ complex (Figure 26B) indicates that putative interactions between the RNA duplex and the Hfq3 proximal side could be transient.

So far, the Hfq distal side has been found to present ARN-rich regions to Crc partner-molecules in the translation-repression assemblies studied here. However, in the Hfq-Crc-*rbsB*$_{110}$ complex, a short RNRN-motif is presented by the Hfq1 distal side to the Hfq2 proximal side (Figure 28B and C, Figure 27, Hfq1b segment in bottom sequence). In particular, basic residues on the Hfq2 N-terminal alpha-helix form hydrogen bonds with phosphate groups in the *rbsB*$_{110}$ backbone (Lys17 and Arg19). In addition, the exposed cytosine base (C84) forms a hydrogen bond with the Hfq2 Arg16 carboxyl group and is stacked between Hfq2 Arg16 and Arg19 (Figure 28C). This interaction between the Hfq proximal side and an RNA substrate has been observed previously in a crystal structure of *E. coli* Hfq and the sRNA RydC (pdbID 4V2S). Here too, basic residues on the N-terminal alpha helix of a Hfq protomer form hydrogen bonds with the RydC backbone (Dimastrogiovanni *et al.*, 2014). Intuitively, this 'out of the box' presentation of an ARN-motif to a Hfq proximal side could help drive higher-order assembly formation by bringing together the Hfq1 and Hfq2 components.

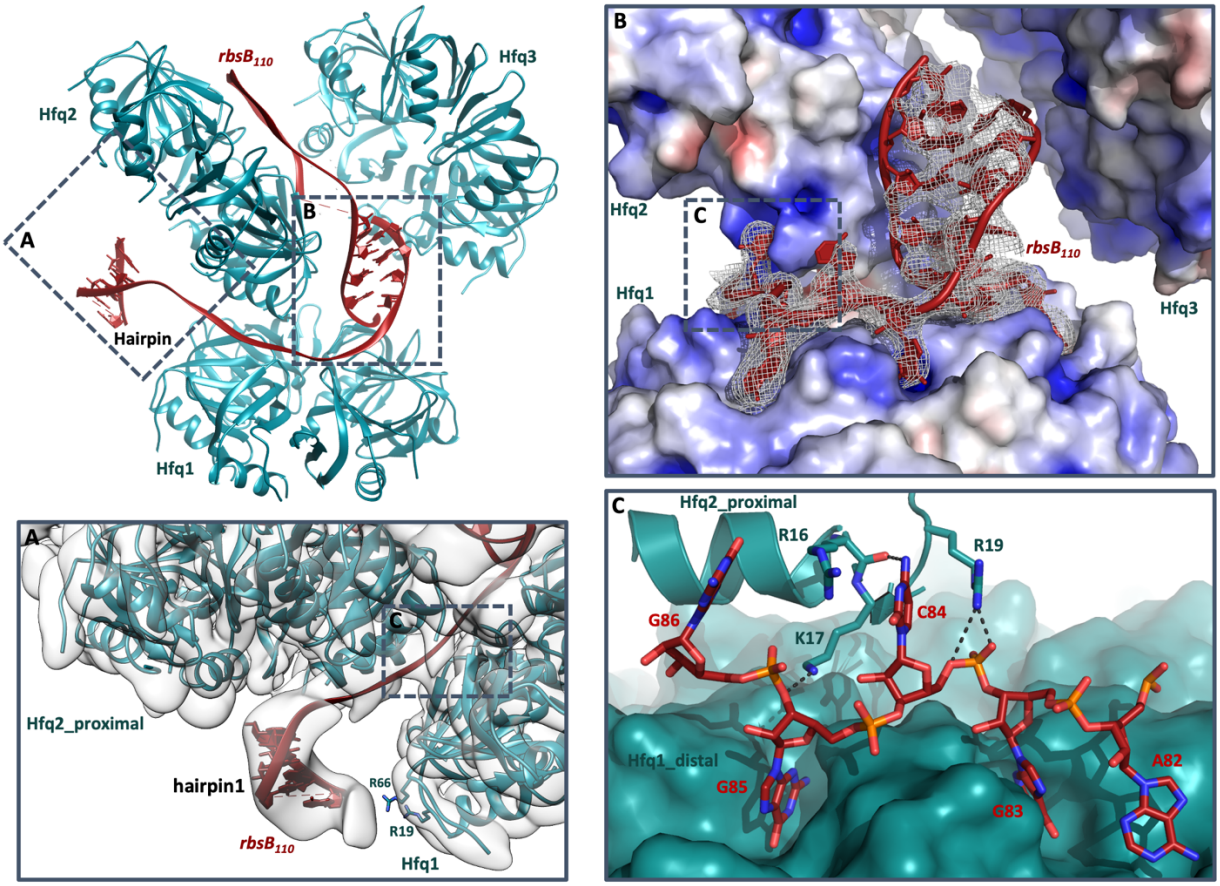


Figure 28: ***Overview of rbsB*$_{110}$ *secondary structures and RNA binding on the Hfq distal side*.** As a reference structure the Hfq-*rbsB*$_{110}$ *sub-assembly is shown, without bound Crc molecules.* ***(A)*** *Density for a short hairpin-loop was observed at the Hfq2 proximal side. Arg66 and Arg19 of the Hfq1 termini are in close proximity to the cryo-EM density for the hairpin-loop.* ***(B)***

Electrostatic presentation of the Hfq surfaces surrounding the rbsB$_{110}$ *duplex. Although the overall environment is basic, there are no obvious basic patches on the Hfq2 rim or Hfq3 proximal side that are close to the central RNA duplex. The cryo-EM map for the* rbsB$_{110}$ *duplex is presented as a grey mesh.* ***(C)*** *The Hfq1 distal side presents the rbsB RNA to Arg16, Lys17 and Arg19 of an N-terminal alpha-helix on the Hfq2 proximal side. These basic residues form hydrogen bonds with the rbsB$_{110}$ phosphate backbone. Surface charges were calculated with the APBS plugin in PYMOL. The map in panel A was filtered with LAFTER (Ramlaul et al., 2020).*

3.4.6.2 Permissive Crc dimerization interfaces help drive higher order assembly

As mentioned above, the quaternary structure of the Hfq-Crc-*rbsB$_{110}$* complex is significantly different from the Hfq-Crc-*amiE$_{105}$* assembly. However, both translation-repression complexes share common architectural features. Crc binds the presented target RNA via the same preferred binding mode in both assemblies. As such, one Hfq-Crc-Crc core unit is formed in the rbsB$_{110}$ assembly, identical to the core units in the Hfq-Crc-a*miE$_{105}$* complex (Figure 29A and Figure 20B). In addition, only one Crc dimer is formed in the *rbsB$_{110}$* assembly, between Crc1 and Crc2 (Figure 29B). These Crc molecules utilise a different dimer interface from the two dimer interfaces observed in the *amiE$_{105}$* complex. In the *rbsB$_{110}$* complex the Crc dimer is asymmetric, where each Crc protomer uses a different binding surface (Figure 29C, dimer 3). So far three different Crc dimers have been observed in the translation-repression complexes described in section 3.4 (Figure 29C).

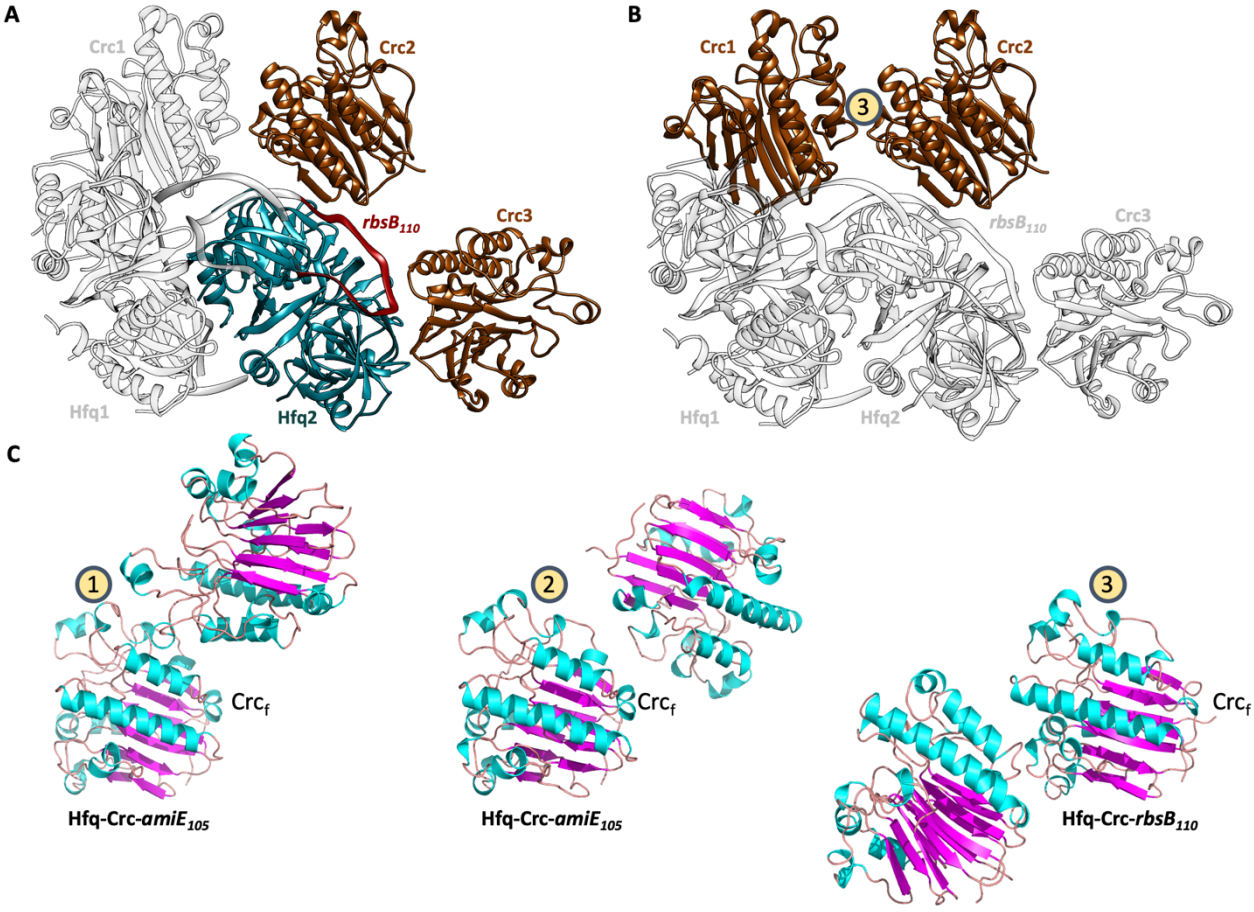


Figure 29: ***A conserved Hfq-Crc-Crc core unit and a novel Crc dimerization interface. (A)*** *A Hfq-Crc-Crc core unit forms the basis for the Hfq-Crc-rbsB$_{110}$ complex, i.e. Crc2 and Crc3 use the same preferred RNA binding mode to engage the presented Brennan crown fold as in the Hfq-Crc-amiE$_{105}$ assembly (RNA binding mode 1). Crc1 also interacts with the presented ARN-segment via*

the same binding mode. ***(B)*** *A single Crc dimer is observed in the $rbsB_{110}$ translation-repression assembly, formed between Crc1 and Crc2.* ***(C)*** *Overview of the Crc dimers observed so far in translation-repression assemblies. Crc dimers 1 and 2 were observed in the $amiE_{105}$ assembly and are symmetric dimers (left and middle). Crc dimer 3 is observed in the $rbsB_{110}$ complex and uses a different dimer interface. Moreover, Crc dimer 3 is an asymmetric dimer, using different patches on each of both molecules, in a stacking-like fashion (right). For comparison, each dimer is depicted relative to a reference Crc, annotated as Crc_f, which is fixed in space for all three dimers. Hfq3 is not depicted in panels A and B for interpretability.*

Table 1: ***Cryo-EM data collection and refinement statistics for the Hfq/Crc/RNA structures.***

Structure	Hfq-Crc-*amiE*$_{6ARN}$ 2:2:2	Hfq-Crc-*amiE*$_{6ARN}$ 2:3:2	Hfq-Crc-*amiE*$_{6ARN}$ 2:4:2	Hfq-Crc-*amiE*$_{105}$	Hfq-Crc-*rbsB*$_{110}$
PDB code	6O1K	6O1L	6O1M	XX	XX
EMDB code	0590	0591	0592	XX	XX
Data collection					
Microscope	FEI Titan Krios G2	FEI Titan Krios G2	FEI Titan Krios G2	FEI Titan Krios G2	FEI Titan Krios G2
Voltage (kV)	300	300	300	300	300
Detector	FEI Falcon III	FEI Falcon III	FEI Falcon III	Gatan K3	Gatan K3
Nominal magnification	92 000	92 000	92 000	105 000	105 000
Pixel size (Å)	1.09	1.09	1.09	0.83	0.83
Electron dose, per frame (e^- /Å^2)	0.37	0.37	0.37	1.03	1.03
Defocus range (μm)	-1.25 / -3	-1.25 / -3	-1.25 / -3	-1.1 / -2.5	-1.1 / -2.5
Exposure (s)	60	60	60	1.9	1.9
Number of micrographs	1308	1308	1308	14996	6674
Reconstruction					
Software	RELION-3.0	RELION-3.0	RELION-3.0	cryoSPARC 2.15	cryoSPARC 2.15
Number of particles used	73 187	45 311	35 376	70 572	148739
Final resolution, $FSC_{0.143}$ (Å)	3.1	3.4	3.2	3.6	3.8
Map-sharpening B factor (Å^2)	-97	-100	-82	-100	-117
Model					
Composition (Hfq:Crc:RNA)	2:2:2	2:3:2	2:4:2	3:4:1	3:3:1
Non-hydrogen atoms	11466	13647	15770	20109	17218
Protein residues	1292	1458	1820	2269	1979
RNA nucleotides	18	18	18	83	52
Molar Mass (kDa)	192	222	252	317	288
Refinement					
Software	Refmac5/Phenix/ Isolde	Refmac5/Phenix/ Isolde	Refmac5/Phenix/ Isolde	Refmac5/Phenix/ Isolde	Refmac5/Phenix/ Isolde

Correlation coefficient, masked	0.79	0.77	0.76	0.81	0.75
$FSC_{0.5}$ (model-map)	3.3	3.8	3.6	3.9	3.6
Validation					
MolProbity score	1.64	1.77	1.69	1.67	1.50
Clash score	5.07	5.44	5.26	5.06	2.80
Ramachandran					
Favoured, overall (%)	94.57	92.64	93.81	94.02	95.71
Allowed, overall (%)	5.27	7.30	6.19	5.75	3.67
Outlier, overall (%)	0.16	0.06	0	0.22	0.63
R.m.s. deviations					
Bond length (Å)	0.004	0.007	0.005	0.015	0.015
Bond angle (°)	0.84	0.88	0.908	1.8	1.9

4 Discussion

Although functional studies have elucidated the cooperation of global post-transcriptional regulators such as Hfq and Crc to coordinate the fate of transcripts, there has been a lack of high-resolution structural data that further explores these orchestrated cellular processes. This chapter fills that gap by describing for the first time how a global and conserved RNA chaperone (Hfq) presents the wide translation initiation region of mRNA targets to an effector molecule (Crc) to govern direct translational repression in the form of diverse higher order, multicomponent assemblies. As such, Hfq and Crc constitute a highly cooperative molecular switch via which *Pseudomonas* can rapidly adapt to changing environments, adding to its well-known intrinsic robustness as a pathogenic superbug.

4.1 Origins and purpose of the Brennan crown in the Hfq-Crc-*amiE* complex

Both Hfq and Crc are known for controlling the translation of mRNAs subject to in CCR (Sonnleitner and Bläsi, 2014; Moreno *et al*., 2015; Kambara *et al.*, 2018). The primary focus of this chapter has been the structural reconnaissance of the translational repression of the mRNA *amiE*, encoding aliphatic amidase. The primary target sequence in *amiE* is the RNA translation initiation site, an A-rich fragment at its 5'-end. Upon binding, the RNA translation initiation site occupies almost entirely the distal side of the Hfq and the similarities with the crystal structure of a poly-A_{18} fragment in complex with *E. coli* Hfq are striking (Link *et al.*, 2009) (Figure 7A and B). The latter added greatly to understanding how RNAs bind to RNA chaperones like Hfq and revealed the stringent rules for recognition of A-R-N type motifs by the Hfq distal side in *E. coli* (A is adenine, R is a purine, and N is any base). Schulz *et al.* proposed that the exposed bases (at the Entry/Exit site, 'N' in the A-R-N motif) could mediate RNA to RNA interactions (Schulz *et al.*, 2017). The structures in this chapter, however, show how the exposed bases and RNA backbone are presented for recognition by Crc to form a cooperative assembly that mediates translational repression of *amiE* when the preferred carbon source is available. Interestingly, A-R-N recognition by the Hfq distal side is not supported in Gram-positive bacteria such as *Staphylococcus aureus* and *Bacillus subtilis* (Horstmann *et al.*, 2012) (Figure 7C). However, this does not mean that Gram-positive bacteria don't exploit translational control mechanisms similar to that of *Pseudomonas* Hfq-Crc assemblies. Instead, the recognition of target RNAs by Hfq is likely to be governed by different rules and partner proteins in these organisms.

The first event in Hfq-Crc mediated regulation is the recognition of ARN-rich regions in the target RNA by Hfq (Figure 30). It seems puzzling that Hfq can combine a low nanomolar affinity for target RNAs with a disproportionally high RNA binding turnover in the cell and a correspondingly fast cellular response (Fender, 2010). The K_D values measured *in vitro* are in the sub-nanomolar range, which would indicate binding half-lives of well over 1 hour. In contrast, the cellular responses facilitated by Hfq are on a timescale of 1-2 minutes, suggesting that nascent RNAs are rapidly cycled through the cellular Hfq pool (Fender *et al.*, 2010; Wagner, 2013; Santiago-Frangos & Woodson; 2018). To reconcile these two observations, it has been suggested that RNA bound by Hfq can be displaced by competitors from the cellular RNA pool in a stepwise process which was coined as active cycling (Fender *et al.*, 2010; Wagner, 2013). An entry point for understanding how active cycling of competing RNAs on the

Hfq distal side works is provided by Molecular Dynamics (MD) simulations. In particular, MD simulations of the Brennan crown fold of *amiE*$_{6ARN}$ reveal a 'dynamic recognition' mechanism whereby the A-site bases can flip back and forth between *syn*- and *anti*-conformations, without change of interactions in the A-site pocket (Figure 10). The rapid transitions between these two conformations likely expose the binding pockets to other, competing RNAs, which can swiftly replace a candidate RNA. Interestingly, binding of Crc strongly attenuates the *syn*-/*anti* flipping behaviour in the simulations, thereby stabilising the interface by prolonging the Hfq-RNA interaction.

Taken together, the molecular dynamics simulations provide insights into the active cycling mechanisms utilised by the Hfq distal side in the earliest stages of Hfq-Crc assembly formation by describing a dynamic interface for RNA binding. Future NMR experiments on short ARN-triplet repeat regions bound to Hfq would test the predictions of the MD simulations for a mixed population of *syn/anti*-conformations for the A-site nucleobases. Lastly, as the MD simulation studies only involved RNA bound to the distal side of Hfq, the 'dynamic recognition' mechanism alone may not be the ultimate answer to the *in vivo* affinity-response discrepancy.

4.2 'On the fly' assembly of Hfq-Crc-RNA complexes

The question that remained after obtaining these initial Hfq-Crc-*amiE*$_{6ARN}$ structures (Pei *et al.*, 2019) was how Hfq and Crc assemble on the whole 5'-end of the *amiE* transcript to form a translation-repression complex. In this context it should be noted that several mRNA targets that are directly repressed by Hfq and Crc, including *amiE*, contain multiple putative Hfq binding motifs upstream and downstream of the start codon (Sonnleitner *et al*., 2018) (Figure 11).

Mobility shift assays with longer 5' fragments of *amiE* and other mRNA targets predicted to form translation-repression complexes illustrate how higher order assembly formation depends on the molar ratios of the components (Figure 12). A broad cryo-EM analysis presented in Figure 22 shows what some of the lower order intermediates of a repressor complex might look like for the *amiE*$_{105}$ target. Whether such intermediate assemblies are formed in the cell remains to be answered. However, Kambara *et al.* (2018) propose a mechanism whereby Hfq chaperones bind nascent transcripts as they spool from the RNA polymerase. These RNA chaperones then present the RNA to Crc as Brennan folds (Figure 30). 3D reconstructions of two intermediate Hfq-Crc-*amiE*$_{105}$ complexes illustrate how co-transcriptional formation of the translation-repression complex could happen 'on the fly' (Figure 23, Figure 30). Interestingly, a recent study on bacterial ribosomal RNA synthesis visualises how other RNA polymerase super complexes can form 'on the fly' during transcription (Huang *et al.*, 2020). In particular, Huang *et al.* showed how during rRNA synthesis, a composite RNA chaperone assembles around the RNA-exit tunnel and promotes transcription completion and co-transcriptional RNA folding. It is plausible that Hfq and Crc assemble near the RNA Polymerase exit tunnel as well.

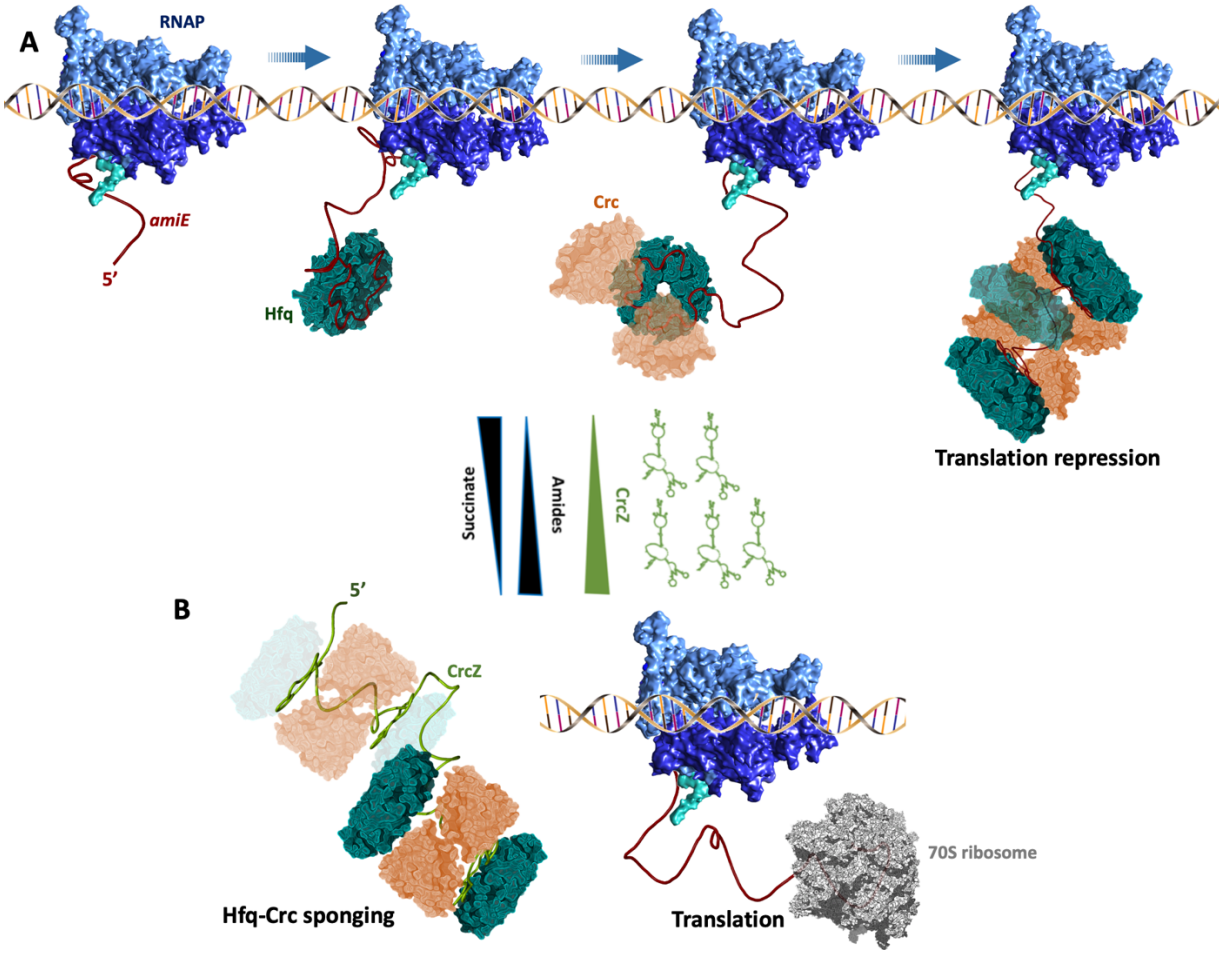


Figure 30: ***Schematic of carbon catabolite repression regulation in Pseudomonas cells. (A)*** *When succinate is abundant as a nutrient, the 5'-end ARN-repeat region of a nascent amiE transcript (red) is recognised by Hfq and presented to a pair of Crc molecules. Additional Hfq and Crc molecules can then participate to form a full repressor complex, which prevents translation of amiE: CCR is in place (Carbon Catabolite Repression).* ***(B)*** *When succinate is exhausted, CrcZ (green) expression rises. CrcZ is rich in ARN-repeat regions and can sequester Hfq and Crc away from target mRNAs, such as amiE. Translation of amiE can now occur, allowing the cell to metabolise other carbon sources: CCR is relieved. RNAP, RNA Polymerase.*

4.3 Hfq and Crc enable quaternary structure variability in repressor complexes

The observation that Crc has no RNA binding capacity on its own and does not form homodimers in solution indicates that the components in translation-repression complexes interact mutually through chelate cooperative effects (Milojevic *et al.*, 2013). The composition and quaternary structure of a full mRNA repressor complex will ultimately depend on the ARN-repeat content in the broad 5' region of a target mRNA. For *amiE*, the full repressor complex consists of three Hfq chaperones, presenting the RNA to four Crc partners (Figure 31, left). All three Hfq molecules present ARN-rich segments of *amiE* to the Crc components via their distal side. Interestingly, secondary structure elements in the *amiE* 5'-end region are recognised by other faces of the participating Hfq components. For example, a short hairpin-loop element is bound by the proximal side of the central Hfq hexamer (Figure 19B). A similar interaction was observed recently for a Hfq-DNA crystal structure by the Brennan group (Orans *et al.*, 2020) (Figure 19C). It is likely that such secondary structure elements also help define the eventual quaternary structure of an effective translation-repression assembly.

Given the sequence dependency for RNA presentation by the Hfq distal side, other mRNA targets, encoding different ARN-contents, will result in alternative repressor assemblies. This is illustrated by a cryo-EM model of the translation-repression complex formed on $rbsB_{110}$, the 5'-end region of a different Hfq-Crc target, *rbsB* (Figure 31, right). Three Hfq chaperones participate in the assembly but have different relative orientations when compared to the *amiE* repressor complex. Three Crc molecules then bind the presented ARN-rich motifs to further shield the RNA from ribosomes. Secondary structure elements in the $rbsB_{110}$ sequence are recognised by Hfq proximal sides, such as a short hairpin loop at the 3'-end (Figure 28A). In addition, a peculiar sheet-like duplex was observed at the centre of the assembly (Figure 28B).

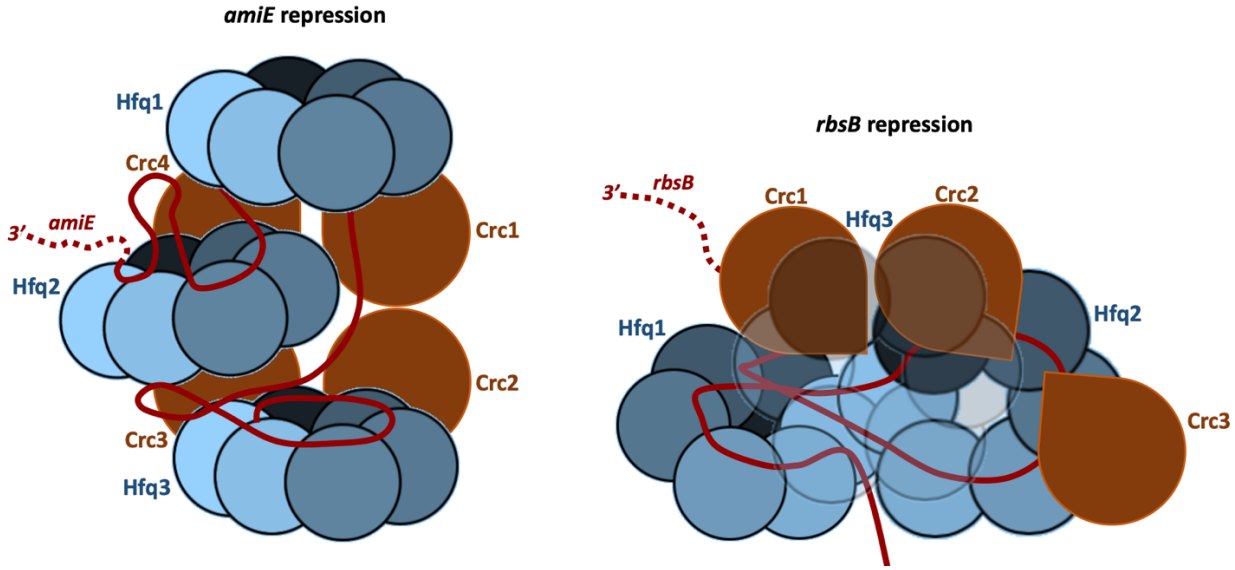


Figure 31: ***Schematic presentation of the amiE and rbsB translation-repression assemblies.*** *Three Hfq chaperones present amiE to four Crc molecules (left). On the other hand, two Hfq hexamers present rbsB to three Crc molecules, with a third Hfq (Hfq3) loosely binding to the assembly (right).*

Although the quaternary structures of the *rbsB* and *amiE* assemblies are entirely different (Figure 31), the lower order building blocks are the same. Specifically, one or more Hfq-Crc-Crc core elements are recognisable in each assembly (Figure 20A & B, and Figure 29A). The question arises as to how two small effector proteins cope with the enormous sequence variability present in the pool of more than 100 target RNAs to effectively form many different repressor complexes in the cell. Part of the answer likely lies in the multiple permissive dimer interfaces harboured by Crc. In the $amiE_{105}$ and $rbsB_{110}$ assemblies alone, three different Crc dimers are observed (Figure 29C). This hard-coded self-interaction polymorphism allows Crc to participate in structurally diverse Hfq-RNA configurations and drive their higher order assembly. In addition, Hfq-Crc repressor complexes appear to form via many small and distant interaction surfaces, creating opportunities for quaternary structure variability. The latter also makes that the complexes are conformationally heterogeneous (Figure 15 and Figure 26). Lastly, Hfq is a homo-hexameric chaperone of which each protomer is rich in RNA binding patches. The ability to employ such a collection of RNA binding modes allows Hfq to 'fit' in many different RNA-mediated quaternary structures. For example, in the structures above Hfq was shown to present RNA not only to Crc but also to a secondary Hfq's distal side ($amiE_{105}$ assembly, Figure 19A) or proximal side ($rbsB_{110}$ assembly, Figure 28C), which drives formation of

higher order assemblies. In summary, the cryo-EM studies in this chapter describe the architectural principles that govern an extensive translation-control mechanism specific to the *Pseudomonadaceae* family.

4.4 Relief of translation-repression by an ARN-rich sRNA

The research presented in this chapter mainly focuses on the architectural features of the Hfq-Crc repressor complex. How translational repression is relieved/prevented in the cell, has not yet been discussed. Pull-down assays have shown that both Hfq and Crc form a complex with CrcZ, a 426nt-long non-coding RNA (Moreno *et al.*, 2015; Sonnleitner *et al.* 2018). The CrcZ sRNA has several ARN-rich stretches which could serve as potential sites for Hfq presentation to Crc (Sonnleitner & Bläsi, 2014). In the presence of less preferred carbon sources, e.g. amides, CrcZ expression levels increase (Sonnleitner *et al.*, 2009) and the RNA sequesters Hfq and Crc away from mRNA targets (Figure 30), thereby relieving translational repression of catabolic genes. This enables expression of genes encoding proteins required for metabolising other carbon sources. How the CrcZ RNA displaces target mRNAs from Hfq-Crc or in turn gets displaced by these mRNAs for changing carbon availability remains unknown. However, the assemblies are likely to be dynamic and the displacement process might resemble that proposed for the stepwise exchange of sRNAs on Hfq (Fender et al., 2010). The latter could happen via the putative 'dynamic recognition' mechanisms described in section 3.3.1. Cryo-EM experiments to elucidate the quaternary structure of Hfq-Crc-CrcZ complexes have been initiated.

4.5 Targeting translation-repression complexes in *Pseudomonas*

Prokaryotic transcription-translation coupling has recently gained significant impact with structural studies of bacterial expressomes *in vitro* and *in situ* (Kohler *et al.* 2017, O'Reilly *et al.*, 2020, Webster *et al.* 2020). The Hfq-Crc regulon forms an alternative gene-regulation route that breaks transcription-translation coupling in *Pseudomonads* by sequestering the ribosome binding region in the mRNA. As mentioned above, the regulatory spectrum of Hfq and Crc was recently found to be much broader than initially expected, with Hfq binding more than 600 nascent transcripts co-transcriptionally, more than 100 of which in concert with Crc (Kambara *et al.*, 2018). These findings showed that Hfq and Crc together regulate gene expression beyond catabolite repression. Insights into post-transcriptional regulation of gene expression in pathogens such as *P. aeruginosa* may provide good candidate targets for novel drug design. In this regard, the Hfq-Crc molecular switch is a good starting point, as they have been shown to control key metabolic and virulence processes in *Pseudomonas* species (O'Toole *et al.*, 2000; Sonnleitner *et al.*, 2003; Sonnleitner *et al.*, 2006; Linares *et al.*, 2010; Huang *et al.*, 2012; Zhang *et al.*, 2012; Zhang *et al.*, 2013; Sonnleitner and Bläsi, 2014; Pusic *et al.*, 2016). Disrupting Crc recognition of the Brennan fold on the Hfq distal side could counter metabolic regulation and virulence specifically in Pseudomonads. A recent study, for example, showed how overproduction of the aliphatic amidase AmiE strongly reduced biofilm formation and almost fully attenuated virulence in a mouse model of acute lung infection (Clamens *et al.*, 2017). Novel drugs that specifically counteract repression of *amiE* translation could induce the phenotype described by Clamens *et al*. (2017). In summary, the high-resolution structures presented here provide a good template for novel approaches to interfere with carbon metabolism and virulence in a multidrug resistant pathogen.

5 Materials and methods

5.1 Hfq synthesis and purification

A Hfq-deficient *E. coli* strain bearing the pKEHfqPae plasmid (Bläsi group, MFPL, University of Vienna) was grown in a 50 ml overnight LB culture supplemented with 0.2% wt/v glucose, 15 µg/ml kanamycin, 50 µg/ml ampicillin. 4 L of LB (with the same additives) were then inoculated with 5 ml of the preculture and grown to OD_{600nm} 0.6, followed by induction with 1mM IPTG (final concentration). The cells were harvested after 4 hours of shaking at 37°C by centrifugation at 5000 g for 20 minutes at 4°C. Finally, the cells were resuspended in 20-50ml lysis buffer (50mM Tris-HCl pH 8, 1.5 M NaCl, 250 mM $MgCl_2$, 1mM β-mercaptoethanol, 1 mM EDTA, 1mM PMSF) and frozen at -80°C.

Thawed Cells were supplemented with 20 µg/ml DNase I and lysed using an Avestin Emulsiflex C5 homogeniser (5 passes, 1000 bars). The lysate was spun down at 35000 g for 20 minutes at 4°C, after which the supernatant was heated to 85°C in a water bath for 45 minutes. The precipitate was removed by a 20000 g spin at 4°C for 15 minutes and 1 M $(NH_4)_2SO_4$ was gradually added to the supernatant. The precipitate was again spun down at 20000 g for 15 minutes (4°C), then the supernatant was filtered with a 0.42 µm Sartorius Minisart syringe filter. The sample was applied to a 5ml HiTrap Butyl HP column (GE Lifesciences) equilibrated in Buffer butyl-A (50 mM Tris-HCl pH 8, 1.5 M NaCl, 1.5 M $(NH_4)_2SO_4$, 0.5 mM β-mercaptoethanol, 0.5 mM EDTA and 0.1 mM PMSF). After loading, 10 column volumes of buffer A were used to wash off contaminants, and a 0-100% linear gradient of buffer butyl-B (50 mM Tris-HCl pH8, 200 mM NaCl, 0.5 mM β-mercaptoethanol, 0.5 mM EDTA and 0.1 mM PMSF) was used to elute Hfq. The eluted protein was diluted twofold in buffer Hep-A (50 mM Tris-HCl pH 8, 100 mM NaCl, 0.5 mM β-mercaptoethanol, 0.5 mM EDTA and 0.1 mM PMSF) and loaded on a 5ml HiTrap Heparin column (GE Lifesciences) equilibrated with buffer Hep-A. 10 column volumes of Buffer Hep-A where then used to wash of any remaining contaminants. Hfq was eluted with a linear 0-60% gradient of buffer Hep-B (50 mM Tris-HCl pH 8, 2 M NaCl). Next, the peak fractions were pooled and concentrated in an Amicon Ultra centrifugal filter unit (10 kDa cutoff) to a final volume of 500 µl. The sample was then loaded on a Superdex 200 Increase 10/300 GL (GE Lifesciences) equilibrated with Buffer SEC-A (50mM Tris-HCl pH 7.5, 200 mM NaCl, 10% v/v glycerol). Peak fractions were flash frozen and stored at -80°C. An SDS-PAGE denaturing gel was run for the peak fractions to assess quality in purity (Figure 3A).

5.2 Crc synthesis and purification

pETM14lic-6His-Crc was transformed into competent *E. coli* BL21DE3 cells via standard heat shock transformation and cells were plated on LB-agar plates supplemented with 50 µg/ml kanamycin. A preculture of the cells was grown overnight in 50 ml LB medium supplemented with 50 µg/ml kanamycin. 4x800 ml of LB, supplemented with 50 µg/ml kanamycin, 0.2% glucose and 2 mM Mg_2SO_4, were inoculated with 4 ml of the preculture. 3 mM of IPTG (final concentration) was used to induce expression at OD_{600nm} of 0.6. After 3 hours of shaking the cells were

harvested by centrifugation at 5020 g for 20 minutes. The pellet was resuspended in 50 ml Ni-A buffer (50 mM Tris-HCl pH 8, 300 mM NaCl, 10 mM imidazole, 1mM β-mercaptoethanol, 0.1 mM PMSF).

The thawed Cells were supplemented with 20 µg/ml DNase I and 20 µg/ml RNase A and lysed using an Avestin Emulsiflex C5 homogeniser (5 passes, 1000 bars). The lysate was spun down for 30 minutes at 30000 g (4°C), and the supernatant was loaded on a 5ml HiTrap chelating column coated with $NiSO_4$ and equilibrated in buffer Ni-A. The column was washed with 10 column volumes of buffer Ni-W (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, pH 8) and eluted with a linear 0-60% gradient of buffer Ni-B (50 mM NaH_2PO_4, 300 mM NaCl, 500 mM imidazole, pH 8). Peak fractions were pooled and dialyzed in 50 mM Hepes pH 8, 150 mM NaCl, 1 mM β-mercaptoethanol and the concentration was measured with a NanoDrop spectrophotometer (Thermo Fisher). For each mg of protein, 20 µg PreScission Protease (Sigma Aldrich) was added to cleave the His-tag. After 2 hours of incubation at 4°C, the sample was run over a nickel column to remove the cleaved His-tags and PreScission Protease and the flow through was concentrated with an Amicon Ultra centrifugal filter (5 kDa cutoff). The sample was loaded on a Superdex 200 Increase 10/300 GL (GE Lifesciences) equilibrated with Buffer SEC-A (50 mM Hepes pH 8, 150 mM NaCl, 1 mM TCEP (tris(2-carboxyethyl) phosphine) and 10% v/v glycerol). Peak fractions were flash frozen and stored at -80°C. An SDS-PAGE denaturing gel was run for the peak fractions to assess purity (Figure 3B).

5.3 Electrophoretic mobility shift assays (EMSA)

For the EMSAs, 4 µM stocks of Hfq and Crc were prepared in binding-buffer (20 mM Tris-HCl pH 8, 40 mM NaCl, 10 mM KCl, 1 mM $MgCl_2$), and a 2 µM stock of the RNAs was prepared in miliQ water (RNase free). RNAs were annealed at 50°C for 3 minutes before adding the proteins. 4% poly-acrylamide (PAA) gels were used to study complex formation (6.73 ml acrylamide:Bis-acrylamide, 5 ml 10x TBE, 37.7 ml miliQ water, 500 µl APS and 50 µl TEMED). Hfq was titrated in a mixture with selected RNAs ($amiE_{105}$, $rbsB_{110}$, $ptxS_{112}$, $acsA_{70}$ and $estA_{119}$ were all provided by the Bläsi group, MFPL, University of Vienna) at different ratios in the presence or absence of an excess of Crc (Figure 11). The RNA concentration was kept constant at 200 nM. After 15 minutes of incubation at 37°C, samples were mixed with an equal volume of loading buffer (50% v/v glycerol, 50% v/v Binding buffer, 5 mM DTT), prior to loading them on the gel. Gels were run at 4°C in 1x TBE running buffer and stained with SYBR gold (Figure 11).

5.4 Cryo-EM sample preparation

The synthetic 18-mer $amiE_{6ARN}$ RNA (5´-AAAAAUAACAACAAGAGG-3´) used in these studies consists of six tripartite binding motifs (Sonnleitner and Bläsi, 2014) and was chemically synthesized. The $amiE_{6ARN}$ RNA was annealed by heating it at 95°C for 5 min followed by 50°C for 10 min and 37°C for 10 min. The RNA was incubated with the Hfq hexamer at a 1:1 molar ratio on ice for 20 min to form a Hfq- $amiE_{6ARN}$ complex. Then an equimolar ratio of Crc was added. The mixture was incubated on ice for 30 min prior to fractionation by size exclusion chromatography using a Superdex 200 Increase 10/300 GL column equilibrated in Binding-buffer (20 mM HEPES, pH 7.9, 10 mM KCl, 40 mM NaCl, 1 mM $MgCl_2$, and 2 mM TCEP). The peak fractions were pooled, and the sample was split in two. Half the sample was cross-linked with bis(sulfosuccinimidyl)suberate (BS3) at 150 µM for 30 min on ice, followed by

quenching at 37.5 mM Tris-HCl pH 8.0. The other half of the sample was not crosslinked. Initial samples were prepared with the help of Dr. Xue Pei (Luisi group, Department of Biochemistry, University of Cambridge) and imaged by Dr. Shaoxia Chen (MRC-LMB, Cambridge).

For Hfq-Crc suppressor assembly on *amiE*$_{105}$, which was provided by the Bläsi group (MFPL, University of Vienna), a simpler but equally effective sample preparation approach was used. The RNA was annealed at 50°C for 3 minutes. Hfq and Crc were mixed at different ratios prior to addition of *amiE*$_{105}$ (800 nM final concentration). After incubation on ice for 1 hour, the mixture was diluted 7-fold right before loading onto grids. The Hfq-Crc-*rbsB*$_{110}$ repressor complex was prepared following a similar procedure. *rbsB*$_{110}$ was annealed at 50°C for 3 minutes. Hfq/Crc were mixed at 2.8 µM and 9 µM, respectively, after which the RNA was added at 400 nM. After incubation at room temperature for 15 minutes and on ice for 1 hour, the mixture was diluted 4-fold prior to grid preparation.

5.5 Grid preparation

Graphene oxide (GO) grids were prepared from regular Quantifoil r1.2/1.3 grids. A 2mg/ml graphene oxide dispersion (Sigma Aldrich) was diluted 10-fold and spun down at 300 g for 30 seconds to remove big aggregates. The dispersion was then diluted another 10-fold before applying 1 µl to glow-discharged grids (0.29 mbar, 15 mA, 2 minutes, Pelco Easiglow glow discharger). After drying out, grids were stored in a grid box for 24-48 hours prior to usage. 3 µl of sample was applied to the GO grids and after 30s of incubation, excess sample was blotted away and frozen in liquid ethane to achieve a thin layer of vitreous ice (blot force -4 to 0, blot time 3 seconds, Vitrobot markIV (FEI)). Grids were screened on a 200 kV Talos Arctica (FEI) (Cryo-EM facility, department of Biochemistry, University of Cambridge) and movies recorded on a 300 kV Titan Krios (FEI) with either a Falcon III (FEI) or K3 (Gatan) direct electron detector (MRC-LMB and BioCem facility, Department of Biochemistry, University of Cambridge).

5.6 Single particle analysis, model building and refinement

All datasets were processed in Warp, CryoSparc and Relion3 (Scheres, 2012; Punjani *et al.*, 2017; Tegunov & Cramer, 2019). For the structures with *amiE*$_{6ARN}$ (section 3.2), whole frame motion correction was performed with motioncorr2 followed by CTF estimation using gctf (Zhang, 2016; Zheng et al., 2017) within Relion-3.0 (Scheres, 2012). Final resolution estimates were calculated after the application of a soft binary mask and phase randomisation and determined based on the gold standard FSC = 0.143 criterion (GS-FSC) (Scheres and Chen, 2012; Chen et al., 2013). For the BS3 crosslinked complex, manually picked particles were used to generate suitable references for autopicking. An initial model was generated using an SGD algorithm (Punjani et al., 2017). During 3D classification, three different complexes could be resolved: 2Hfq:2Crc:2amiE$_{6ARN}$ (2:2:2), 2Hfq:3Crc:2amiE$_{6ARN}$ (2:3:2) and 2Hfq:4Crc:2amiE$_{6ARN}$ (2:4:2). To properly separate, validate and refine the three classes, the same 3D classification was rerun with the new 2:3:2 model as reference model, lowpass filtered to 20 Å resolution. C2 symmetry was observed and imposed for the 2:2:2 and 2:4:2 complexes. Each of the classes was then refined to sub-3.5 Å resolution, followed by per-particle frame alignment for movement correction and per-frame damage weighting in a particle polishing step. The resulting 'polished' particles were subjected to a final refinement round

with a 3D mask and solvent flattening. All reference models were lowpass filtered to 60 Å prior to refinement. The dominant class (2:2:2) had a resolution of 3.1 Å (GS-FSC) (Table 1). Local resolution calculations were done with the Relion local resolution estimation tool. The 2:4:2 and 2:3:2 maps were sharpened locally with LocScale and LocalDeblur to better resolve the additional Crc components (Jakobi et al., 2017; Ramírez-aportela et al., 2018).

Datasets from samples with $amiE_{105}$ and $rbsB_{110}$ were pre-processed on the fly with Warp (Tegunov and Cramer, 2019). Particle sets were cleaned up in CryoSparc (Punjani *et al.*, 2017) via repetitive 2D classifications and heterogeneous refinements. Further, extensive classifications in 2D were used to classify the different assemblies observed on the grid (Figure 14). For those assemblies that had enough diverse projections, high resolution maps were generated with non-uniform refinement in cryoSPARC (Punjani *et al.*, 2019) and global and per particle CTF refinements (Table 1). The Hfq -2Crc-$amiE_{105}$ (147K particles), 2Hfq-3Crc-$amiE_{105}$ (99K particles) and 3Hfq-4Crc-$amiE_{105}$ (18K particles) assemblies were refined to 3.2Å, 3.9Å and 3.6Å respectively. The Hfq-Crc-$rbsB_{110}$ map was refined to 3.8Å.

The crystal structures for *P. aeruginosa* Crc (PDB code 4JG3) and Hfq (PDB code 1U1T) were manually docked into the EM density map as rigid bodies in Chimera (Pettersen et al., 2004). The $amiE_{6ARN}$, $amiE_{105}$ and $rbsB_{110}$ sequences were manually built into the density using Coot (Emsley et al., 2010). Refmac5 and Phenix real-space refinement with global energy minimization, NCS-restraints, group B-factor and geometry restraints were used to iteratively refine the multi-subunit complexes at high resolution, followed by manual corrections for Ramachandran and geometric outliers in Coot and ISOLDE (Table 1) (Emsley et al., 2010; Murshudov et al., 2011; Afonine et al., 2012; Croll, 2018). Model quality was evaluated with MolProbity (Williams et al., 2018). For the $amiE_{6ARN}$ models, in silico 2 Å maps were generated from the atomic models and FSC validation against the experimental maps was performed with the EMDB Fourier shell correlation server (EMBL-EBI). For the $amiE_{105}$ and $rbsB_{110}$ structures model to map FSC were calculated by Refmac5 as part of the ccpem suite (Burnley *et al.*, 2017).

5.7 *In vivo* translational repression assays with an *amiE::lacZ* reporter gene

These experiments were performed by Dr. Elisabeth Sonnleitner at the Bläsi group and are described in Pei *et al.*, 2019.

5.8 Molecular dynamics simulations of the Hfq-*amiE* interface

These experiments were performed in collaboration with Dr. Miroslav Krepl and Prof. Jiri Sponer ((CEITEC, Brno, Czech Republic). Details are described in Krepl, Dendooven *et al.* (submitted for publication).

6 References

Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., and Adams, P.D. (2012). Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. D Biol. Crystallogr. *68*, 352-367.

Bandyra, K.J., Sinha, D., Syrjanen, J., Luisi, B.F., and De Lay, N.R. (2016). The ribonuclease polynucleotide phosphorylase can interact with small regulatory RNAs in both protective and degradative modes. RNA *22*, 360–372.

Burnley, T., Palmer, C. M., & Winn, M. (2017). Recent developments in the CCP-EM software suite. *Acta crystallographica. Section D, Structural biology*, *73*(Pt 6), 469–477. https://doi.org/10.1107/S2059798317007859

Chen, S., McMullan, G., Faruqi, A.R., Murshudov, G.N., Short, J.M., Scheres, S.H.W., and Henderson, R. (2013). High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryo-microscopy. Ultramicroscopy *135*, 24–35.

Clamens, T., Rosay, T., Crépin, A., Grandjean, T., Kentache, T., Hardouin, J., Bortolotti, P., Neidig, A., Mooij, M., Hillion, M., Vieillard, J., Cosette, P., Overhage, J., O'Gara, F., Bouffartigues, E., Dufour, A., Chevalier, S., Guery, B., Cornelis, P., Feuilloley, M. and Lesouhaitier, O. (2017). The aliphatic amidase AmiE is involved in regulation of Pseudomonas aeruginosa virulence. *Scientific Reports*, 7(1).

Croll, T.I. (2018). ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Cryst. D74*

Dimastrogiovanni, D., Fröhlich, K. S., Bandyra, K. J., Bruce, H. A., Hohensee, S., Vogel, J., & Luisi, B. F. (2014). Recognition of the small regulatory RNA RydC by the bacterial Hfq protein. *eLife*, *3*, e05375. https://doi.org/10.7554/eLife.05375

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486-501.

Fender, A., Elf, J., Hampel, K., Zimmermann, B., and Wagner, E.G. (2010). RNAs actively cycle on the Sm-like protein Hfq. Genes Dev. *24*, 2621-2626.

Fernandez, L., Breidenstein, E.B., Taylor, P.K., Bains, M., de la Fuente-Nunez, C., Fang, Y., Foster, L.J., and Hancock, R.E. (2016). Interconnection of post-transcriptional regulation: the RNA-binding protein Hfq is a novel target of the Lon protease in *Pseudomonas aeruginosa*. Sci. Rep. *6*, 26811.

Heitzinger, D.A. (2016) Impact of Hfq on antibiotic susceptibility of *Pseudomonas aeruginosa*. Master Thesis, University of Vienna.

Holloway, B.W., Krishnapillai, V., and Morgan, A.F. (1979). Chromosomal genetics of *Pseudomonas*. Microbiol. Rev. *43*, 73-102.

Horstmann, N., Orans, J., Valentin-Hansen, P., Shelburne, S.A., and Brennan, R.G. (2012) Structural mechanism of *Staphylococcus aureus* Hfq binding to an RNA A-tract. Nucleic Acids Res. *40*, 11023-11035.

Huang, J., Sonnleitner, E., Ren, B., Xu, Y., and Haas, D. (2012). Catabolite repression control of pyocyanin biosynthesis at an intersection of primary and secondary metabolism in *Pseudomonas aeruginosa*. Appl. Environ. Microbiol. *78*, 5016-5020.

Huang, Y. H., Hilal, T., Loll, B., Bürger, J., Mielke, T., Böttcher, C., Said, N., & Wahl, M. C. (2020). Structure-Based Mechanisms of a Molecular RNA Polymerase/Chaperone Machine Required for Ribosome Biosynthesis. *Molecular cell*, S1097-2765(20)30553-0. Advance online publication. https://doi.org/10.1016/j.molcel.2020.08.010

Ikeda, Y., Yagi, M., Morita, T., and Aiba, H. (2011). Hfq binding at RhlB-recognition region of RNase E is crucial for the rapid degradation of target mRNAs mediated by sRNAs in *Escherichia coli.* Mol. Microbiol. *79*, 419-432.

Jakobi, A.J., Wilmanns, M. & Sachse, C. 2017. Model-based local density sharpening of cryo-EM maps. *Elife*. 6:1–26

Kambara, T.K., Ramsey, K.M., and Dove, S.L. (2018). Pervasive targeting of nascent transcripts by Hfq. Cell Reports *23*, 1543-1552.

Kohler, R., Mooney, R. A., Mills, D. J., Landick, R., & Cramer, P. (2017). Architecture of a transcribing-translating expressome. Science, *356(6334)*, 194–197. https://doi.org/10.1126/science.aal3059

Lehtinen, J., Munkberg, J., Hasselgren, J., Laine, S., Karras, T., Aittala, M. & Aila, T.. (2018). Noise2Noise: Learning Image Restoration without Clean Data. *Proceedings of the 35th International Conference on Machine Learning, in PMLR*80:2965-2974

Link, T.M., Valentin-Hansen, P., and Brennan, R.G. (2009). Structure of *Escherichia coli* Hfq bound to polyriboadenylate RNA. Proc. Natl. Acad. Sci. U.S.A. *106*, 19292-19297.

Linares, J.F., Moreno, R., Fajardo, A., Martínez-Solano, L., Escalante, R, Rojo, F., and Martínez, J.L. (2010). The global regulator Crc modulates metabolism, susceptibility to antibiotics and virulence in *Pseudomonas aeruginosa.* Environ. Microbiol. *12,* 3196-3212.

Lu, P., Wang, Y., Zhang, Y., Hu, Y., Thompson, K.M., and Chen, S. (2016). RpoS-dependent sRNA RgsA regulates Fis and AcpP in *Pseudomonas aeruginosa.* Mol. Microbiol. *102*, 244–259.

Melamed, S., Peer, A., Faigenbaum-Romm, R., Gatt, Y. E., Reiss, N., Bar, A., Altuvia, Y., Argaman, L., & Margalit, H. (2016). Global Mapping of Small RNA-Target Interactions in Bacteria. *Molecular cell*, *63*(5), 884–897. https://doi.org/10.1016/j.molcel.2016.07.026

Melamed, S., Adams, P. P., Zhang, A., Zhang, H., & Storz, G. (2020). RNA-RNA Interactomes of ProQ and Hfq Reveal Overlapping and Competing Roles. *Molecular cell*, *77*(2), 411–425.e7. https://doi.org/10.1016/j.molcel.2019.10.022

Miller, J.H. (1972). Experiments in Molecular Genetics. (New York: Cold Spring Harbor Press)

Milojevic, T., Grishkovskaya, I., Sonnleitner, E., Djinovic-Carugo, K., and Bläsi U. (2013). The *Pseudomonas aeruginosa* catabolite repression control protein Crc is devoid of RNA binding activity. PLoS One *8,* e64609.

Mohanty, B.K., Maples, V.F., and Kushner, S.R. (2004). The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in *Escherichia coli.* Mol. Microbiol. *54,* 905–920.

Moreno, R., Hernandez-Arranz, S., La Rosa, R., Yuste, L., Madhushani, A., Shingler, V., and Rojo,F. (2015). The Crc and Hfq proteins of *Pseudomonas putida* cooperate in catabolite repression and formation of ribonucleic acid complexes with specific target motifs. Environ. Microbiol*. 17*, 105-118.

Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Vagin, A.A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. D Biol. Crystallogr. *67*, 355–367.

Orans, J., Kovach, A. R., Hoff, K. E., Horstmann, N. M., & Brennan, R. G. (2020). Crystal structure of an Escherichia coli Hfq Core (residues 2-69)-DNA complex reveals multifunctional nucleic acid binding sites. *Nucleic acids research*, *48*(7), 3987–3997. https://doi.org/10.1093/nar/gkaa149

O'Reilly, F.J., Xue, L., Graziadei, A., Sinn, L., Lenz, S., Tegunov, D., Blötz, C., Hagen, W.J.H., Cramer, P., Stülke, J., Mahamid, J. & Rappsilber, J. (2020). In-cell architecture of an actively transcribing-translating expressome. bioRxiv, doi:10.1101/2020.02.28.970111.

O'Toole, G.A., Gibbs, K.A., Hager, P.W., Phibbs, P.V.Jr., and Kolter, R. (2000). The global carbon metabolism regulator Crc is a component of a signal transduction pathway required for biofilm development by *Pseudomonas aeruginosa*. J. Bacteriol. *182*, 425-431.

Panja, S., Schu, D.J., and Woodson, S.A. (2013). Conserved arginines on the rim of Hfq catalyze base pair formation and exchange. Nucleic Acids Res. *41*, 7536-7546.

Pantelic, R.S., Meyer, J.C., Kaiser, U., Baumeister, W., and Plitzko, J.M. (2010). Graphene oxide: a substrate for optimizing preparations of frozen-hydrated sample. J. Structural Biol. *170*, 152-156.

Pei, X.Y.*, Dendooven, T.*, Sonnleitner, E., Chen, S., Bläsi U, Luisi, B.F. (2019) Architectural principles for Hfq/Crc-mediated regulation of gene expression. eLife;8:e43158.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera - a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605-1612.

Punjuani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods *14*, 290-296.

Punjani A., Zhang H., Fleet D.J. (2019) Non-uniform refinement: Adaptive regularization improves single particle cryo-EM reconstruction. bioRxiv, doi:10.1101/2019.12.15.877092.

Punjani, A. & Fleet, D.J. (2020). 3D Variability Analysis: Directly resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM images. *bioRxiv*. doi:10.1101/2020.04.08.032466

Pusic, P., Tata, M., Wolfinger, M.T., Sonnleitner, E., Häussler, S. and Bläsi, U. (2016). Cross-regulation by CrcZ RNA controls anoxic biofilm formation in *Pseudomonas aeruginosa*. Sci. Rep. *6*, 39621.

Ramírez-aportela, E., Vilas, J.L., Melero, R., Conesa, P. *et al.* 2018. Automatic local resolution-based sharpening of cryo-EM maps. bioRxiv, pp. 1–21. https://doi.org/10.1101/433284

Ramlaul, K., Palmer, C.M., Aylett, C.H.S. (2020) Mitigating Local Over-fitting During Single Particle Reconstruction with SIDESPLITTER. *bioRxiv*. January 2020:2019.12.12.874081. doi:10.1101/2019.12.12.874081

Resch, A., Večerek, B., Palavra, K. and Bläsi, U. (2010). Requirement of the CsdA DEAD-box helicase for low temperature riboregulation of *rpoS* mRNA. RNA Biol. *7*, 796–802.

Ripin, N., Boudet, J., Duszczyk, M. M., Hinniger, A., Faller, M., Krepl, M., Gadi, A., Schneider, R. J., Šponer, J., Meisner-Kober, N. C., et al., Molecular basis for AU-rich element recognition and dimerization by the HuR C-terminal RRM. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116*, 2935-2944.

Rojo, F. (2010). Carbon catabolite repression in *Pseudomonas*: optimizing metabolic versatility and interactions with the environment. FEMS Microbiol. Rev*. 34*, 658–684.

Santiago-Frangos, A., Jeliazkov, J.R., Gray, J.J., and Woodson, S.A. (2017). Acidic C-terminal domains autoregulate the RNA chaperone Hfq. Elife *6*, e27049.

Santiago-Frangos, A., Kavita, K., Schu, D. J., Gottesman, S., and Woodson, S. A. (2016). C-terminal domain of the RNA chaperone Hfq drives sRNA competition and release of target RNA. Proc. Natl. Acad. Sci. U.S.A. *113*, E6089–E6096.

Santiago-Frangos, A., Fröhlich, K. S., Jeliazkov, J. R., Małecka, E. M., Marino, G., Gray, J. J., Luisi, B. F., Woodson, S. A., & Hardwick, S. W. (2019). *Caulobacter crescentus* Hfq structure reveals a conserved mechanism of RNA annealing regulation. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(22), 10978–10987. https://doi.org/10.1073/pnas.1814428116

Sauer, E., Schmidt, S., and Weichenrieder, O. (2012). Small RNA binding to the lateral surface of Hfq hexamers and structural rearrangements upon mRNA target recognition. Proc. Natl. Acad. Sci. U.S.A. *109*, 9396-9401.

Scheres, S.H. (2012). A Bayesian view on cryo-EM structure determination. J. Mol. Biol. 415, 406-418.

Scheres, S.H., and Chen, S. (2012). Prevention of overfitting in cryo-EM structure determination Nat. Methods *9,* 853–854.

Schulz, E.C., Seiler, M., Zuliani, C., Voight, F., Rybin, V., Pogenberg, V., Mucke, N., Wilmanns, M., Gibson, T.J. and Barabas, O. (2017). Intermolecular base stacking mediates RNA-RNA interaction in a crystal structure of the RNA chaperone Hfq. Sci. Rep. *7*, 9903.

Schumacher, M.A., Pearson, R.F., Møller, T., Valentin-Hansen, P., and Brennan, R.G.(2002). Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein. EMBO J. *21*, 3546-3556.

Sonnleitner, E., Abdou, L., and Haas, D. (2009). Small RNA as global regulator of carbon catabolite repression in *Pseudomonas aeruginosa*. Proc. Natl. Acad. Sci. U.S.A. *106,* 21866–21871.

Sonnleitner, E., and Bläsi, U. (2014). Regulation of Hfq by the RNA CrcZ in *Pseudomonas aeruginosa* carbon catabolite repression. PLoS Genet. *10*, e1004440.

Sonnleitner, E., Hagens, S., Rosenau, F., Wilhelm, S., Habel, A., Jäger, K.E., and Bläsi, U. (2003) Reduced virulence of a *hfq* mutant of *Pseudomonas aeruginosa* O1. Microb. Pathog. *35*, 217–228.

Sonnleitner, E., Schuster, M., Sorger-Domenigg, T., Greenberg, E.P., and Bläsi, U. (2006). Hfq-dependent alterations of the transcriptome profile and effects on quorum sensing in *Pseudomonas aeruginosa*. Mol. Microbiol. *59*, 1542–1558.

Sonnleitner, E., Wulf, A., Campagne, S., Pei, X., Wolfinger, M., Forlani, G., Prindl, K., Abdou, L., Resch, A., Allain, F., Luisi, B., Urlaub, H. and Bläsi, U. (2018). Interplay between the catabolite repression control protein Crc, Hfq and RNA in Hfq-dependent translational regulation in Pseudomonas aeruginosa. Nucleic Acids Res. *46*, 1470-1485.

Sukhodolets, M.V., and Garges, S. (2003). Interaction of *Escherichia coli* RNA polymerase with the ribosomal protein S1 and the Sm-like ATPase Hfq. Biochemistry *42*, 8022–8034.

Swanson, B.L., Colmer, J.A., Hamood, A.N. (1999). The Pseudomonas aeruginosa Exotoxin A Regulatory Gene, ptxS: Evidence for Negative Autoregulation. J Bacteriol *181*, 4890–4895.

Tan, Y. Z., Baldwin, P. R., Davis, J. H., Williamson, J. R., Potter, C. S., Carragher, B., & Lyumkis, D. (2017). Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nature methods*, *14*(8), 793–796. https://doi.org/10.1038/nmeth.4347

Tegunov, D., Cramer, P. (2019). Real-time cryo-electron microscopy data preprocessing with Warp. *Nat Methods* 16, 1146–1152. https://doi.org/10.1038/s41592-019-0580-y

Valentini, M., Garcia-Maurino, S.M., Perez-Martinez, I., Santero, E., Canosa, I., and Lapouge, K. (2014). Hierarchical management of carbon sources is regulated similarly by the CbrA/B systems in *Pseudomonas aeruginosa* and *Pseudomonas putida*. *Microbiology 160***,** 2243-2252.

Van den Bossche, A., Ceyssens, P., De Smet, J., Hendrix, H., Bellon, H., Leimer, N., Wagemans, J., Delattre, A., Cenens, W., Aertsen, A., Landuyt, B., Minakhin, L., Severinov, K., Noben, J. and Lavigne, R. (2014). Systematic Identification of Hypothetical Bacteriophage Proteins Targeting Key Protein Complexes of Pseudomonas aeruginosa. *Journal of Proteome Research 13*, pp.4446-4456.

Večerek, B., Beich-Frandsen, M., Resch, A., and Bläsi, U. (2010). Translational activation of *rpoS* mRNA by the non-coding RNA DsrA and Hfq does not require ribosome binding. Nucleic Acids Res. *38*, 1284–1293.

Vogel, J., and Luisi, B.F. (2011). Hfq and its constellation of RNA. Nat. Rev. Microbiol. *9*, 578-589.

Wagner, E.G., and Romby, P. (2015). Small RNAs in bacteria and archaea: who they are, what they do, and how they do it. Adv. Genet. *90*, 133–208.

Williams, C., Headd, J., Moriarty, N., Prisant, M., Videau, L., Deis, L., Verma, V., Keedy, D., Hintze, B., Chen, V., Jain, S., Lewis, S., Arendall, W., Snoeyink, J., Adams, P., Lovell, S., Richardson, J. and Richardson, D. (2017). MolProbity: More and better reference data for improved all-atom structure validation. *Protein Science 27*, 293-315.

Salvail, H., Caron, M.P., Bélanger, J. and Massé, E. (2013) Antagonistic functions between the RNA chaperone Hfq and an sRNA regulate sensitivity to the antibiotic colicin. EMBO J. *32*, 2764-2778.

Santiago-Frangos, A. and Woodson, S.A. (2018) Hfq Chaperone Brings Speed Dating to Bacterial sRNA. *WIREs RNA*, **9**, e1475.

Večerek B., Moll, I. and Bläsi, U. (2005) Translational autocontrol of the *Escherichia coli hfq* RNA chaperone gene. RNA *11*, 976-984.

Vogel, J., Luisi, B. Hfq and its constellation of RNA. *Nat Rev Microbiol* **9,** 578–589 (2011). https://doi.org/10.1038/nrmicro2615

Wagner, E.G.H. (2013) Cycling of RNAs on Hfq. *RNA Biol.*, **10**, 619-626.

Webster, M.W., Takacs, M., Zhu, C., Vidmar, V., Eduljee, A., Abdelkareem, M., Weixlbaumer, A. (2020). Structural basis of transcription-translation coupling and collision in bacteria. bioRxiv doi:10.1101/2020.03.01.971028.

Wei, Y., Zhang, H., Gao, Z.Q., Xu, J.H., Liu, Q.S. and Dong, Y.H. (2012) Structure analysis of the global metabolic regulator Crc from *Pseudomonas aeruginosa*. IUBMB Life *65*, 50-57.

Worrall, J.A., Gorna, M., Crump, N.T., Phillips, L.G., Tuck, A.C., Price, A.J., Bavro, V.N., and Luisi, B.F. (2008). Reconstitution and analysis of the multienzyme *Escherichia coli* RNA degradosome. J. Mol. Biol. *382*, 870–883.

Yang, N., Ding, S., Chen, F., Zhang, X., Xia, Y., Di, H., Cao, Q., Deng, X., Wu, M., Wong, C., Tian, X., Yang, C., Zhao, J. and Lan, L. (2015). The Crc protein participates in down-regulation of the Lon gene to promote rhamnolipid production and quorum sensing in Pseudomonas aeruginosa. *Molecular Microbiology 96*, 526-547.

Zhang, K. (2016). Gctf: real-time CTF determination and correction. J. Struct. Biol*. 193*, 1-12.

Zhang, L., Chiang, W.C., Gao, Q., Givskov, M., Tolker-Nielsen, T., Yang, L., and Zhang, G. (2012). The catabolite repression control protein Crc plays a role in the development of antimicrobial-tolerant subpopulations in *Pseudomonas aeruginosa* biofilms. Microbiology *158*, 3014–3019.

Zhang, L., Gao, Q., Chen, W., Qin, H., Hengzhuang, W., Chen, Y., Yang, L., and Zhang, G. (2013). Regulation of *pqs* quorum sensing via catabolite repression control in *Pseudomonas aeruginosa*. Microbiology *159*, 1931–1936.

Zheng, S.Q, Palovcak, E., Armache, J.-P., Cheng, Y. and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods *14*, 331-332.

Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., et al. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife*. 7: e42166:1–22

Zuker M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic acids research*, *31*(13), 3406–3415. https://doi.org/10.1093/nar/gkg595

Chapter II: Cooperative PNPase-Hfq carrier complexes protect sRNAs and facilitate handover

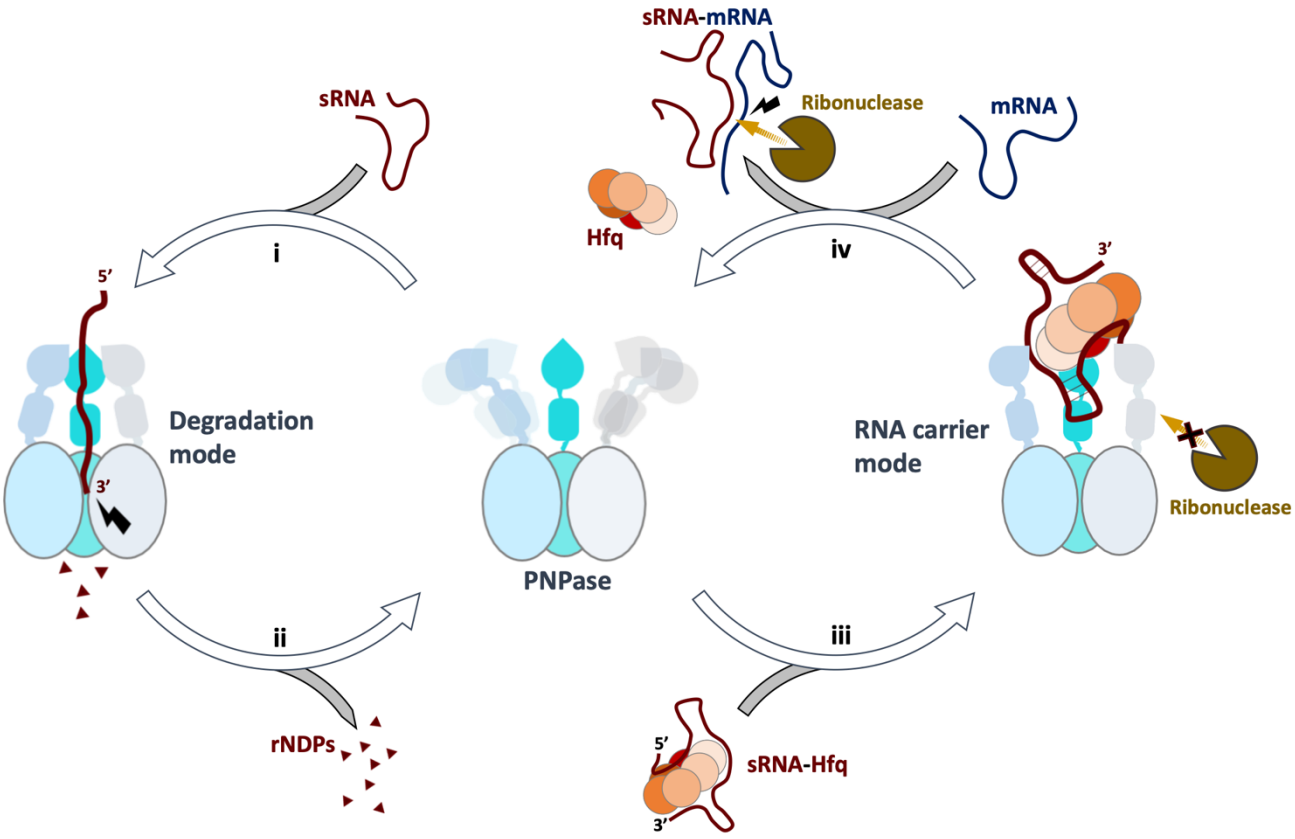


Table of Contents

1 Abstract

Polynucleotide phosphorylase, a conserved exoribonuclease, degrades RNA processively from the 3'-end in a reversible phosphorolytic reaction. Paradoxically, in *Escherichia coli* PNPase can also protect regulatory small RNAs and facilitate their actions *in vivo*, with pleiotropic consequences for gene regulation. This alternative mode of action has been proposed to result from the formation of a stable ternary complex with the RNA chaperone Hfq. In this chapter, cryo-EM structures were solved of PNPase in apo state, in degradation mode and as an sRNA carrier in a ternary complex with Hfq. Comparing these structures reveals how an sRNA is rerouted in the ternary complex through interactions with Hfq and the S1 and KH RNA binding domains of PNPase. The structures rationalise an RNA sequence pattern that favours the ternary assembly and also provide the first visualisation of how a class-II sRNA is presented by Hfq. *In vitro* activity assays show that sRNAs are protected from degradation by other ribonucleases when part of the RNA carrier complex. Gel shift assays further indicate that the sRNA carrier complex can be disbanded by the sRNA target and aid in the sRNA handover process. Finally, *in* vivo reporter assays, performed by collaborators, prove the relevance of the ternary complexes in stabilising sRNAs. Taken together, the PNPase-Hfq carrier complexes protect sRNA species in the cell and potentially facilitate their regulatory actions on target genes.

2 Introduction

In all domains of life, ribonucleases are key players in post-transcriptional regulation of gene expression. These enzymes act through mRNA degradation, rRNA and tRNA maturation, and often impact on the stability of regulatory RNAs. Naturally, ribonucleases must be tightly regulated to prevent uncontrolled destruction of cellular RNA pools (Deutscher, 2015). Polynucleotide phosphorylase (PNPase) (Figure 1A) is an exoribonuclease of ancient evolutionary origin that contributes to RNA degradation and RNA quality control in diverse organisms (Cheng & Deutscher, 2003; Cameron *et al.*, 2018; Danchin, 1997). PNPase was discovered in 1955 by Grunberg-Manago & Ochoa and was shown to catalyse both the phosphorolytic degradation and polymerisation of polynucleotides, depending on inorganic phosphate levels (Grunberg-Manago & Ochoa, 1955). The enzyme was famously utilised to synthesise homo- and hetero-polyribonucleotides to elucidate the genetic triplet code (Lengyel *et al.*, 1961). Decades of subsequent research have established the cellular function of PNPase as catalysis of the 3' to 5' phosphorolysis of substrate RNA (Figure 1A). This reaction requires inorganic phosphate and Mg^{2+} as a cofactor, and releases ribonucleoside diphosphates (rNDPs) as product (Nurmohamed *et al.*, 2009).

In bacteria, PNPase is involved in bulk mRNA turnover and tRNA and rRNA processing (Figure 2). To contribute to steady state RNA turnover PNPase is often abundantly expressed. For example, in *E. coli* and other bacteria it is in the top 5% of most expressed proteins (Wang *et al.*, 2015). Deletion of PNPase in bacteria is known to reduce virulence and increases sensitivity to stressors, which shows its importance for cell survival (Cameron *et al.*, 2018). Although mostly cytosolic, a significant fraction of PNPase (~30% in *E. coli*) associates with the bacterial RNA degradosome in many Gram-negative bacteria. The degradosome is a multicomponent protein complex

responsible for bulk degradation of cellular RNAs using a DEAD-box helicase RhlB, the endonucleolytic activity of RNase E, and the processive exoribonucleolytic activity of PNPase (Plocinski *et al.*, 2019; Cameron *et al.*, 2018; Chandran *et al.*, 2007) (Figure 2). In the last few decades it has become clear that PNPase also plays a key role in the regulation of gene expression via tight control over the small regulatory RNA (sRNA) pool in the cell. Between 50 and 200 nucleotides long, these single-stranded RNAs base-pair with RNA targets and influence their stability or translational efficiency (Waters & Storz, 2009, Dendooven & Luisi, 2017). A well-established mechanism for control of sRNA stability is via degradation by PNPase. Finally, PNPase can also participate in the maturation of CRISPR RNAs involved in phage adaptive immunity (Chou-Zeng & Hatoum-Aslan, 2019; Sesto *et al.*, 2014).

Structural studies have provided insights into the PNPase makeup and how it degrades RNA substrates. PNPase arose through gene duplication of the phosphorolytic enzyme RNase PH, which resulted in the catalytic PH1 and PH2 domains, interlinked by an alpha helical domain (Symmons *et al.*, 2002) (Figure 1A). Notably, the RNase PH core is shared with other evolutionary ancestors, i.e. the eukaryotic and archaeal exosomes. In bacteria these domains assemble co-linearly to form the trimeric, toroidal catalytic core of PNPase, containing a central channel (Figure 1A). The active sites of PNPase reside in the central channel, where the last 3'-end phosphodiester linkage of the RNA substrate is oriented for attack by inorganic phosphate with the assistance of a magnesium cofactor (Nurmohamed *et al.*, 2009, Symmons *et al.*, 2002). Interestingly, only one of two RNase PH domains that constitute the PNPase protomer is catalytically active. The entrance to the central channel of PNPase is guarded by a narrow aperture that is only sufficiently wide to accept single stranded RNA. Ideal substrates for PNPase have a single stranded 3'-end, and polyadenylation of RNAs at this 3'-end may increase their degradation rate by the enzyme. Transcripts containing stem-loop structures, on the other hand, have been shown to be protected *in vivo* (Dar & Sorek, 2018; Spickler & Mackie, 2000). The KH and S1 domains encircle the channel entrance and form three flexibly tethered extensions to the catalytic RNase PH-like core (Figure 1A). The RNA binding patches on these domains are proposed to capture single stranded substrates for degradation in the central channel. Although partial crystal structures have been solved of PNPase constructs, both in apo-form and engaged on a substrate (Nurmohamed *et al.*, 2009, Hardwick *et al.*, 2012), there is little knowledge on how the KH and S1 domains are organised and how they coordinate RNA substrates during degradation. The main hurdle so far has been the inherent flexibility of the KH-S1 portal, leading to poor packing in a crystal lattice.

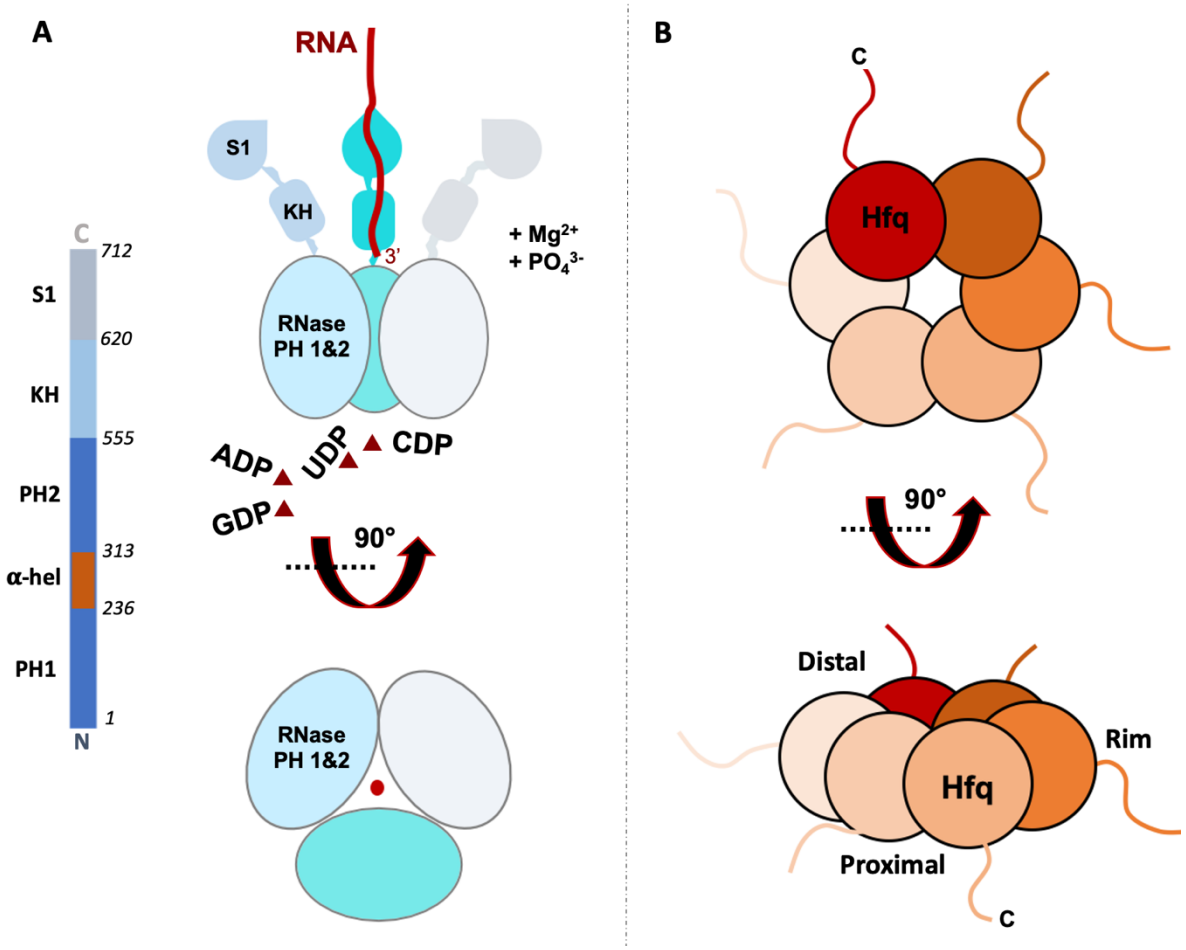


Figure 1: ***Schematic representation of PNPase and Hfq architectures. (A)*** *PNPase forms homo-trimeric assemblies in vivo. The catalytic core of PNPase consists of RNase PH domains (PH1 and PH2), one of which is still active, separated by an α-helical linker (α-hel). Two globular domains, KH and S1, are flexibly tethered to the PNPase core-domain and constitute an RNA entry portal, rich in RNA binding patches.* ***(B)*** *Hfq forms homo-hexameric toroidal multimers with several RNA binding surfaces on its distal-, proximal- and rim sides. The C-terminal tails (indicated by 'C') are natively unstructured and have been shown to help coordinate RNA binding specificity (Santiago-Frangos et al., 2019).*

The research presented in this chapter originates from the peculiar results of a PNPase knock-out experiment (De Lay & Gottesman, 2011). Deletion of the PNPase gene from the *E. coli* or *Salmonella* chromosome stabilises many transcripts, as expected, but paradoxically also results in an increased turnover of many small regulatory RNAs (sRNAs) of both type I and type II classes. This in turn results in a loss of efficiency with which the sRNAs can control the expression of targeted genes (Cameron & De Lay, 2016; De Lay & Gottesman, 2011, Cameron *et al.* 2019). These initial data indicate that PNPase could be a pleiotropic regulator of gene expression, beyond its ribonuclease activities. How PNPase stabilises many sRNAs has been explored by Bandyra *et al.* (2016). In *E. coli*, PNPase has been co-isolated from cells in a complex with Hfq (Figure 1B) (Bandyra *et al.*, 2016). Hfq is highly conserved RNA chaperone and itself a key regulator of many riboregulatory pathways in bacteria with pleiotropic consequences. Notably, Hfq binds to hundreds of small noncoding RNAs and improves their efficacy by aiding base pairing to target mRNAs (Vogel & Luisi, 2011). In solution, Hfq assembles into a ring-like homo-hexamer with RNA binding patches on its proximal, distal, and rim-side (Figure 1B). The C-terminal tail of Hfq is natively unstructured

and has been shown to coordinate RNA binding events (Santiago-Frangos *et al.*, 2016; Santiago-Frangos *et al.*, 2017; Santiago-Frangos *et al.*, 2019). Association with Hfq protects sRNAs from degradation by PNPase and Hfq-sRNA intermediates have been shown to form protective effector complexes with PNPase (Bandyra *et al.*, 2016). These assemblies, called 'RNA carrier' assemblies in this chapter, may even facilitate downstream action of these regulatory RNAs (Figure 2). Previous studies have already shown that the activity of PNPase can be modulated via formation of ribonucleoprotein assemblies or by small ligands (Gatewood & Jones, 2010; Nurmohamed *et al.*, 2011; Tuckerman *et al.*, 2011; Stone *et al.*, 2017). In *Deinococcus radiodurans*, for example, the activity of PNPase on RNA duplexes can be directed through assembly with a noncoding RNA Y-RNA and the Rsr protein. The latter is an ortholog of the eukaryotic Ro protein proposed to act in quality control mechanisms (Chen *et al.*, 2013) and a similar complex may exist in *Salmonella enterica* serovar Typhimurium (Sim & Wolin, 2018).

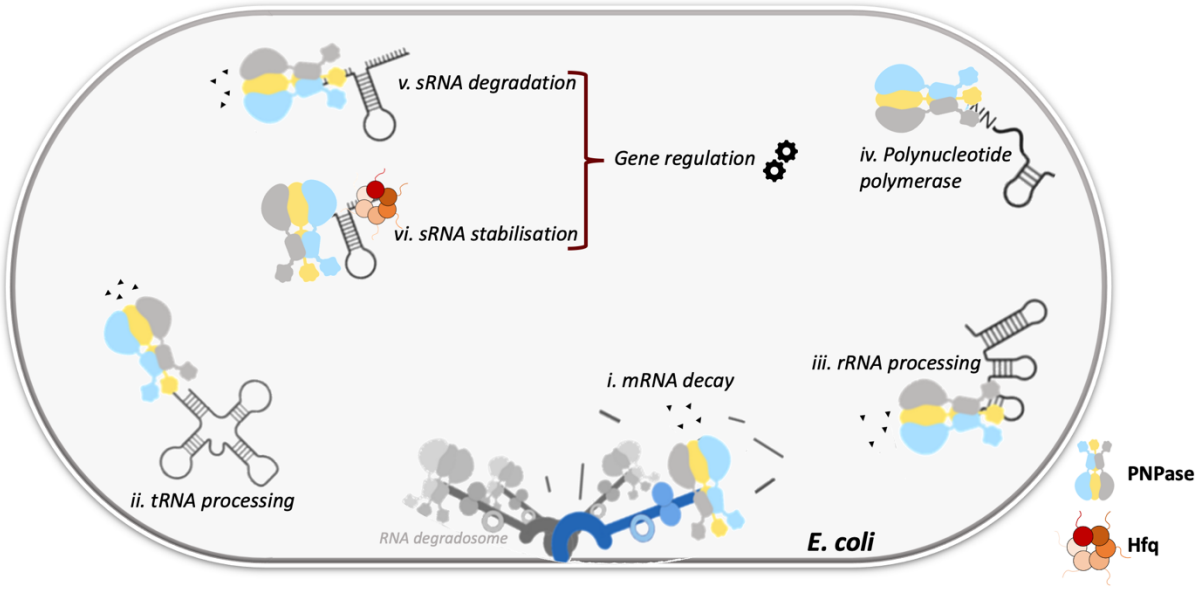


Figure 2: ***Overview of PNPase activity in E. coli.*** *As part of the RNA degradosome, PNPase participates in the bulk degradation of mRNA (i). PNPase is also involved in the maturation of tRNA and rRNA (ii and iii) and has been shown to have polymerase activity at low orthophosphate concentrations (iv). PNPase degrades small regulatory RNA substrates but can form a PNPase-sRNA-Hfq ternary complex in the presence of Hfq (v and vi). The latter has been shown to protect the sRNA substrate rather than degrade it.*

These complex relationships between PNPase and RNA levels in bacteria are equally prevalent in mammalian cells, where PNPase is involved in processes beyond RNA degradation. Mammalian PNPase may play a role in mitochondrial RNA surveillance, degradation and transport of RNA into and out of the organelle (Silva *et al.*, 2018; Cameron *et al.*, 2018; Cheng *et al.*, 2018). In human mitochondria, PNPase interaction with Suv3 helicase enables efficient degradation of redundant RNAs (Cameron *et al.*, 2018) and contributes to surveillance (Silva *et al.*, 2018). Mutations of PNPase result in accumulation of mitochondrial double-stranded RNA that trigger activation of innate immune mechanisms. The localization of PNPase to the mitochondrial inter-membrane space and matrix suggests that the enzyme has a dual role in preventing the formation and release of such double-stranded RNA into the cytoplasm (Dhir *et al.*, 2018). Interestingly, mitochondrial PNPase is also involved in the import of long non-coding RNAs and serves as an RNA carrier protein rather than a ribonuclease. For example, recent data suggest that the

enzyme could help import the RNA component of telomerase in mitochondria for processing. PNPase would also export the processed form back to the cytoplasm (Cheng *et al.*, 2018).

An important structural puzzle is how the mammalian and bacterial exoribonuclease PNPase can govern both ribonucleolytic and carrier modes of action. In this chapter, the aim is to understand how *E. coli* PNPase can switch between these modes of RNA interaction and how the RNA chaperone Hfq facilitates this. Cryo-electron microscopy (cryo-EM) was used to solve a complete structure for *E. coli* PNPase in the apo form, substrate bound form, and in a ternary complex with sRNA substrate and Hfq, i.e. in the RNA carrier form. The apo structure captures the conformational heterogeneity of the KH and S1 domains, and the complex with RNA substrates reveals a conserved path of the RNA to the active site in the RNA-destructive mode. The S1 and KH domains can cooperate to interact with the substrate and facilitate its delivery into the central channel to access the active sites. In the RNA carrier mode, Hfq cooperates with the KH-S1 portal of PNPase to capture the RNA and prevent the 3'-end from entering the central channel, safeguarding it from degradation (Figure 2, vi). The latter assembly also confirms the predicted mode of interaction between small regulatory RNAs of the class-II group and Hfq. These sRNAs are predicted to bind both the proximal and distal sides of Hfq and wrap over the surface of the ring-shaped Hfq hexamer. On the Hfq distal side, a degenerate ARN repeat in the RNA substrate constitutes the interaction surface between Hfq and PNPase, imposing a loose sequence preference. Gel shift assays and activity screens show how the PNPase-Hfq carrier complex protects sRNAs from degradation by other ribonucleases and at the same time can facilitate handover of the sRNA to its RNA target. It is plausible that transport complexes formed by the human mitochondrial PNPase may mimic the mode of action observed for the bacterial ternary complex of Hfq, sRNA and PNPase.

3 Results

3.1 Expression and purification of *E. coli* PNPase and Hfq

Expression and purification of *E. coli* PNPase and Hfq is explained in detail in the materials and methods section (Sections 5.1 and 5.2). Figure 3 shows elution profiles and SDS PAGE gels of PNPase (78 kDa, Figure 3A) and Hfq (11.2 kDa, Figure 3B) after a size exclusion chromatography step. A small tail was observed in the PNPase elution profile (Figure 3A, top), and more contaminants were observed in the corresponding lanes of the SDS-PAGE gel. The Hfq elution profile displays a slightly asymmetric peak (Figure 3B, top). SDS/heat denaturation was not sufficient to fully denature the heat resistant Hfq hexamer (Figure 3B, bottom, Hfq oligomers marked by *). Material from the main peak for both samples are sufficiently pure and homogenous for structural studies.

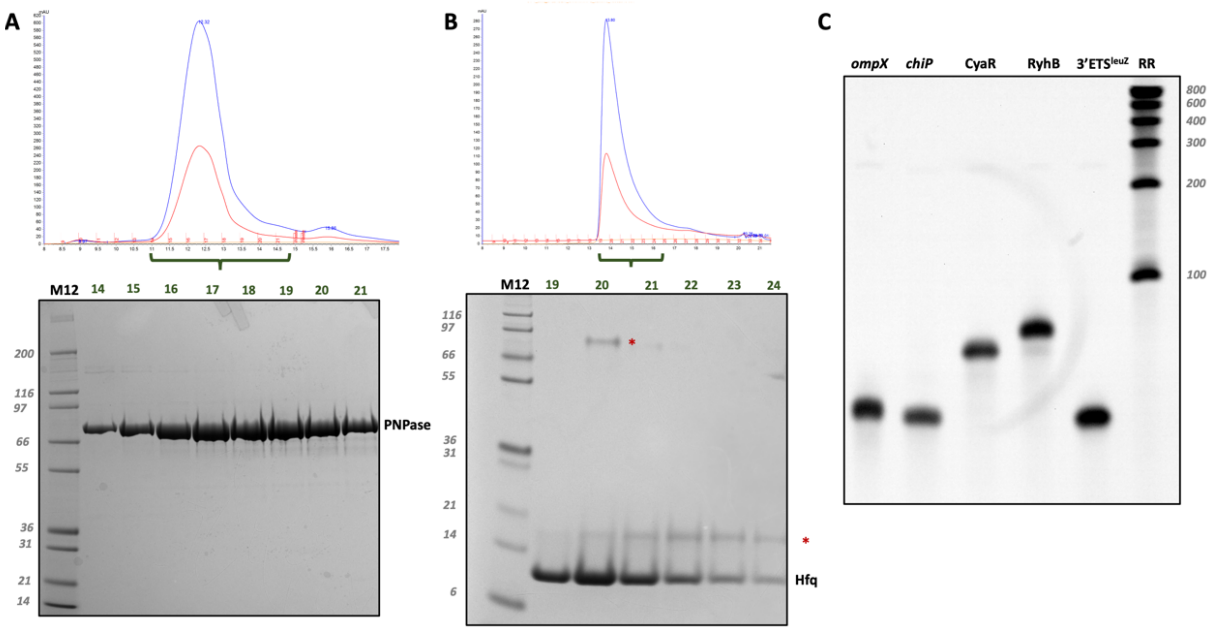


Figure 3: ***Purification of PNPase, Hfq and RNA fragments****. Elution profiles and SDS PAGE gels for PNPase and Hfq after size exclusion chromatography are depicted in* ***(A)*** *and* ***(B)****, respectively. The gels were visualised with Coomassie staining. Both Hfq and Crc are sufficiently pure and homogeneous for further structural studies.* * *Highlights Hfq oligomer bands (A) or contaminants (B).* ***(C)*** *Denaturing PAA/Urea gels were run on purified RNA fragments to assess purity and homogeneity and visualised by fluorescence after SYBR Gold staining. M12, Mark 12 protein standard, RR, Riboruler RNA standard.*

3.2 In vitro transcription and purification of RNA samples

For this project five different RNA fragments, 3'ETS^{leuZ}, RyhB, CyaR, *chiP* and *ompX*, were produced using an in vitro transcription system and purified via electro elution (Materials and Methods section 5.3). Three of these, 3'ETS^{leuZ}, RyhB and CyaR are small regulatory RNAs. The other two, *chiP* and *ompX*, are fragments of selected target mRNAs for RyhB and CyaR respectively, containing the complement that matches the corresponding sRNA seed regions. The biological role for each of these will be described over the next paragraphs. All RNAs produced here were assessed for purity by denaturing gel (Figure 3C). For all experiments described below, the RNA was annealed at 50°C for three minutes prior to mixing it with PNPase, Hfq and/or other RNAs.

3.3 cryo-EM analysis of the apo-PNPase trimer resolves the flexible KH-S1 portal

3.3.1 Preferred orientation at the AWI and CHAPSO

The main goal in this chapter was to explore how Hfq repurposes PNPase from an exoribonuclease into an RNA carrier/chaperone when acting on sRNA substrates. It is likely that the KH-S1 portal is involved in these repurposing events, since these domains are known to be important for capturing RNA substrates (Shi *et al.*, 2008). However, in the published crystal structures for *E. coli* PNPase the KH and S1 domains remain unresolved due to their flexible nature and consequential lack of proper packing in a crystal lattice (Nurmohamed *et al.*, 2009). Given these observations, I set out to solve a more complete structure of *E. coli* PNPase by cryo-EM single particle analysis (SPA), so that a reference 'apo-PNPase' structure would be available to compare with the degradative- and RNA carrier complexes (section 3.4 and section 3.5, respectively).

Samples of the apo-PNPase were iteratively optimized for cryo-EM studies. The main hurdle in getting high resolution reconstructions was preferred orientations of the PNPase particles (Figure 4A-C). This occurs when particles align with the air-water interface during the time between blotting the excess sample away and freezing the grid in liquid ethane (D'Imprima *et al.*, 2019). In this case, 99.99% of all the particles in the imaged grid holes aligned with the air-water interface, which resulted in an overabundance of donut like 'top views' (Figure 4A and B). Only 0.01% of the particles had a different orientation in the vitreous ice (Figure 4C). Subsequent reconstructions were highly anisotropic and showed severe signs of directional smearing, as further evidenced by a disc-like Fourier transform of the map (Figure 4D and E). The apparent absence of the KH-S1 portal, even after significant B-factor blurring, indicates that partial denaturation occurs when PNPase aligns with the air-water interface (Figure 4D).

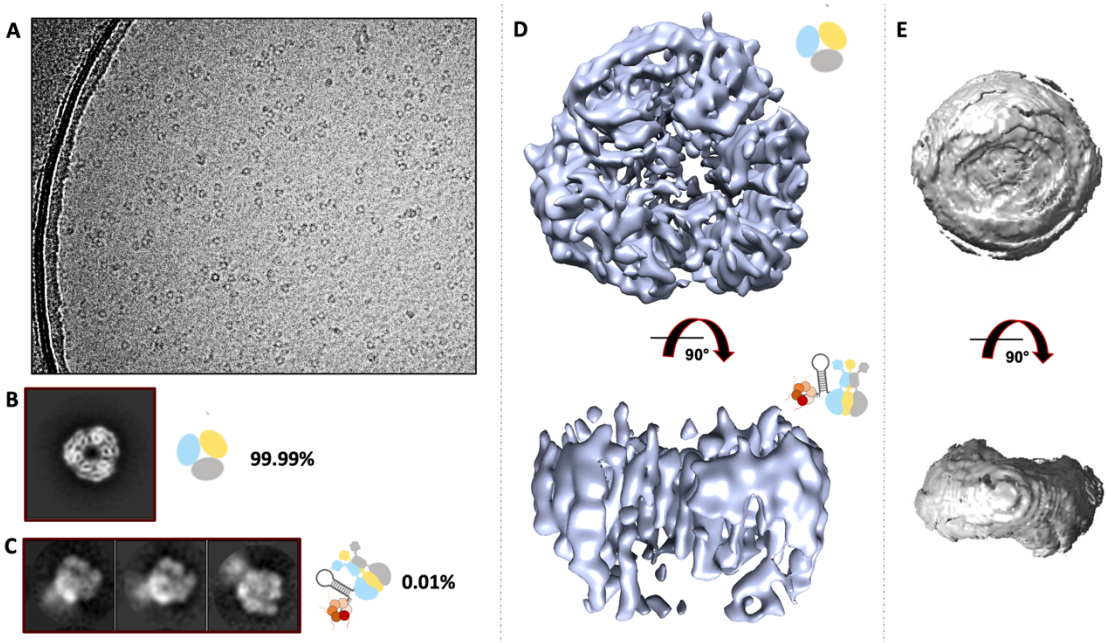


Figure 4: ***Preferred orientation and partial denaturation at the air water interface. (A)*** *Representative cryo-EM image at 3 µm defocus.* ***(B)*** *99.99% of all particles aligned with the air-water interface in a well-defined orientation.* ***(C)*** *Only 0.01% of the picked particles corresponded to side-on views of PNPase.* ***(D)*** *3D reconstructions were smeary and lacked density for the KH-S1*

portal, most likely due to partial denaturation. ***(E)*** *In reciprocal space the map has a disc-like shape, further showing that the 3D reconstruction is highly anisotropic (generated with the 3D-FSC server, Tan et al., 2017).*

To ameliorate preferred orientation artefacts in the reconstructions, several strategies were employed. For example, the cryo-stage was tilted to 45° during data collection (Tan *et al.*, 2017). This resulted in a map that was isotropic up to 7Å (GS-FSC) but did not resolve the KH-S1 portal (data not shown). Therefore, attempts were made to prevent PNPase from denaturing at the air-water interface during grid freezing. Gradual crosslinking with glutaraldehyde combined with ultracentrifugation in a glycerol gradient, a technique called GraFix, only marginally improved the particle-orientation distribution on the grids (Kastner *et al.*, 2007) (data not shown). Spotiton, a piezo electric inkjet dispensing system integrated into a custom designed vitrification robot, was used to reduce the time between sample blotting and plunge freezing of the grid down to 80 ms (Razinkov *et al.*, 2016). Unfortunately, PNPase still aligns with the air water interface in this short time span, and no improvement in particle orientation was observed (data not shown). Finally, detergents such as β-DDM (β-dodecylmaltoside) and Tween20 were added to the PNPase sample at their respective CMC (critical micelle concentration). None of thes additives significantly ameliorated the orientation bias in the vitreous ice (data not shown).

Ultimately, preferred orientation of the PNPase particles was overcome by adding CHAPSO, a zwitterionic detergent, to the sample prior to grid freezing (Figure 5A and B) (Chen *et al.*, 2019). CHAPSO was added at 8 mM, corresponding to the CMC, to a 10 µM PNPase sample seconds before vitrification. Such high protein concentrations are necessary to achieve a suitable number of particles in the grid holes when using CHAPSO as an additive (Figure 5A). Interestingly, the detergent seems to oligomerize into long fibres with low helical rise and small repeating units (Figure 5A and B). Formation of such CHAPSO fibres has not been reported before and a high-resolution 3D reconstruction of these has not been successful so far. 2D class averages and subsequent 3D reconstructions of apo-PNPase in Relion 3.0 (Scheres, 2012; Zivanov *et al.*, 2018) reveal that the particles adopt nearly random orientations in the vitreous ice, resulting in a highly isotropic reconstruction of the PNPase core (Figure 5C-E, Figure 7). An overview of the cryo-EM data processing pipeline is given in Figure 6.

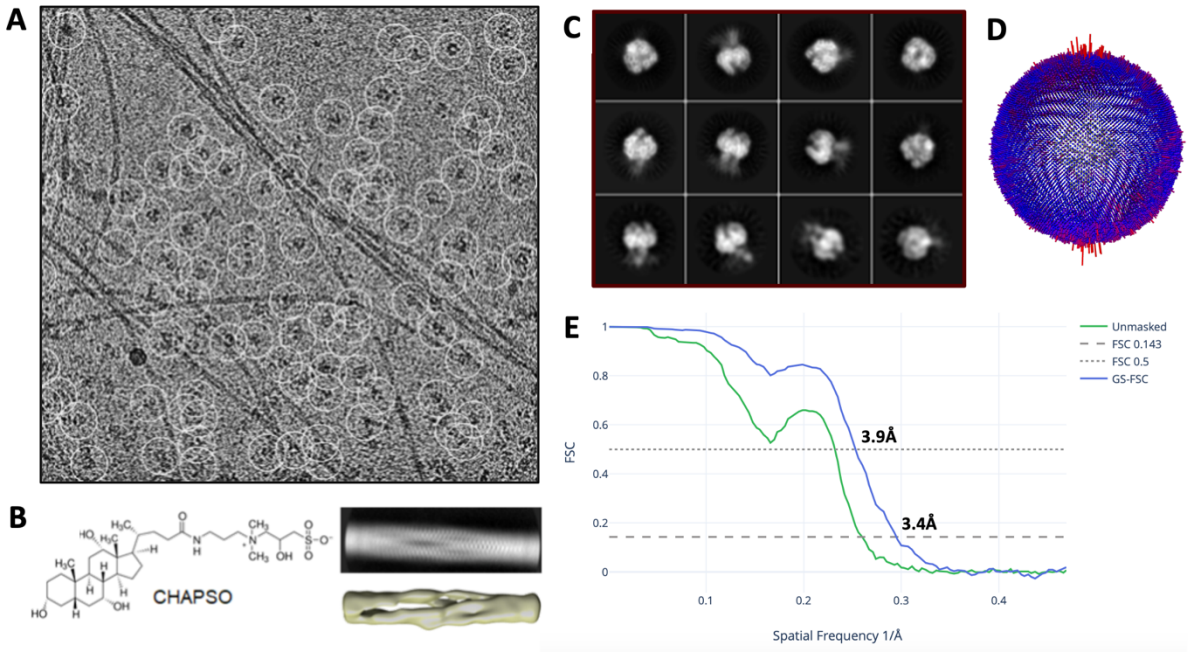


Figure 5: ***CHAPSO randomizes particle orientations in the vitreous ice. (A)*** *Representative cryo-EM image at 3 µm defocus. Individual particles are highlighted by white circles.* ***(B)*** *Chemical structure of CHAPSO, 2D class average and 3D reconstruction of a CHAPSO fibre.* ***(C)*** *2D class averages of the PNPase particles.* ***(D)*** *Angular distributions of the 2D images that contributed to the refined PNPase map, presented as a spherical bar plot, with red bars representing more abundant projections in the images.* ***(E)*** *Global resolution estimation of the refined reconstruction according to the Gold Standard FSC (GS-FSC).*

3.3.2 Symmetry expansion and 3D Variability Analysis of the PNPase KH-S1 portal

Using Relion 3.0 (Scheres, 2012; Zivanov *et al.*, 2018), reconstructions at a global map resolution of 3.4 Å were obtained, with well resolved features for the RNase PH core, but diffuse density for the flexibly tethered KH and S1 domains (Figure 5E and Figure 7A and B, Table 1). During the iterative reconstruction, the symmetric core drives the alignments so that these flexible domains are averaged out in the resulting consensus map. A clean stack of particles was transferred to cryoSPARC for 3D variability analysis (3DVA) of the PNPase particles (Figure 6) (Punjani & Fleet, 2020). After symmetry expansion, 3DVA revealed large molecular motions within the KH-S1 portal (Figure 12A and B). In parallel, symmetry expansion tools in Relion (Zivanov *et al.*, 2018) were used to successfully deconvolute the apo-PNPase protomers and distinguish three distinct conformational sub-states for the KH-S1 domain densities via masked 3D classification (Figure 6 and Figure 7A and C, Materials and methods section 5.7). In these 3D classes, the resolution of the KH domains improved significantly, and the S1 domains, albeit at lower resolution, have stronger density. For docking into the improved cryo-EM map, a complete PNPase homology model was generated with MODELLER (Fiser & Sali, 2003). The *E. coli* PNPase core (Nurmohamed *et al.*, 2009) was extended based on the crystal structure of the complete *Caulobacter crescentus* PNPase (Hardwick *et al.*, 2012) and missing loops were filled in *de novo*. The improvements in the cryo-EM map allowed for local rigid-body docking (Chimera, Pettersen *et al.* 2004) and subsequent molecular dynamics fitting of the KH and S1 domains in the homology model (Namdinator, Kidmose *et al.*, 2019). The observed domain movements correspond to an opening and closing of the KH-S1 portal (Figure 7C). The transition between the states is unlikely to be sterically hindered in the trimeric PNPase assembly (Figure 7B and C). However, given that the KH and S1 domains are

averaged out in consensus alignments, the stable sub-states of each KH-S1 pair are not fully synchronised. It has been shown that the KH and S1 domains are key for recognising and capturing RNA substrates (Stickney *et al.*, 2005). The flexible linkers connecting these domains could help PNPase not only to intercept RNA substrates but could also facilitate different interaction modes between PNPase and substrates or effector molecules in the cell, as will become apparent in the next sections.

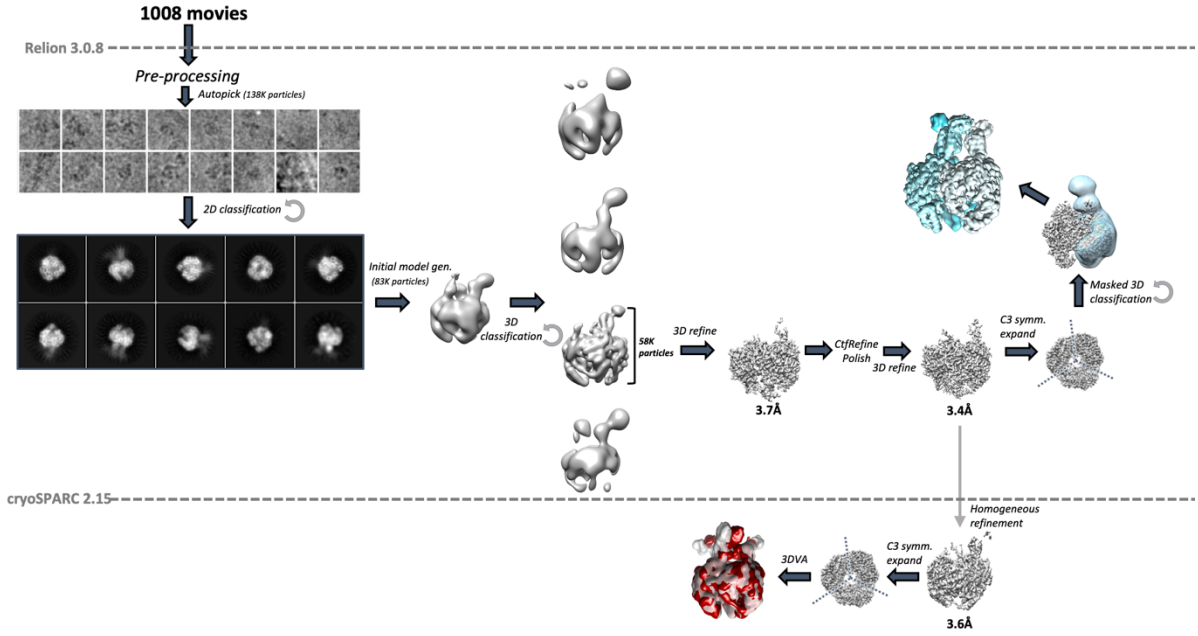


Figure 6: ***Schematic overview of the processing pipeline for apo-PNPase.*** *Pre-processing, which entails motion correction, ctf estimation, manual picking and 2D reference generation was carried out in Relion 3.0.8. Consensus maps were refined to 3.4 Å (GS-FSC) prior to C3 symmetry expansion and focused 3D classifications (Relion). In parallel, a clean particle stack was transferred to cryoSPARC for symmetry expansion and subsequent 3D variability analysis (Punjani & Fleet, 2020).*

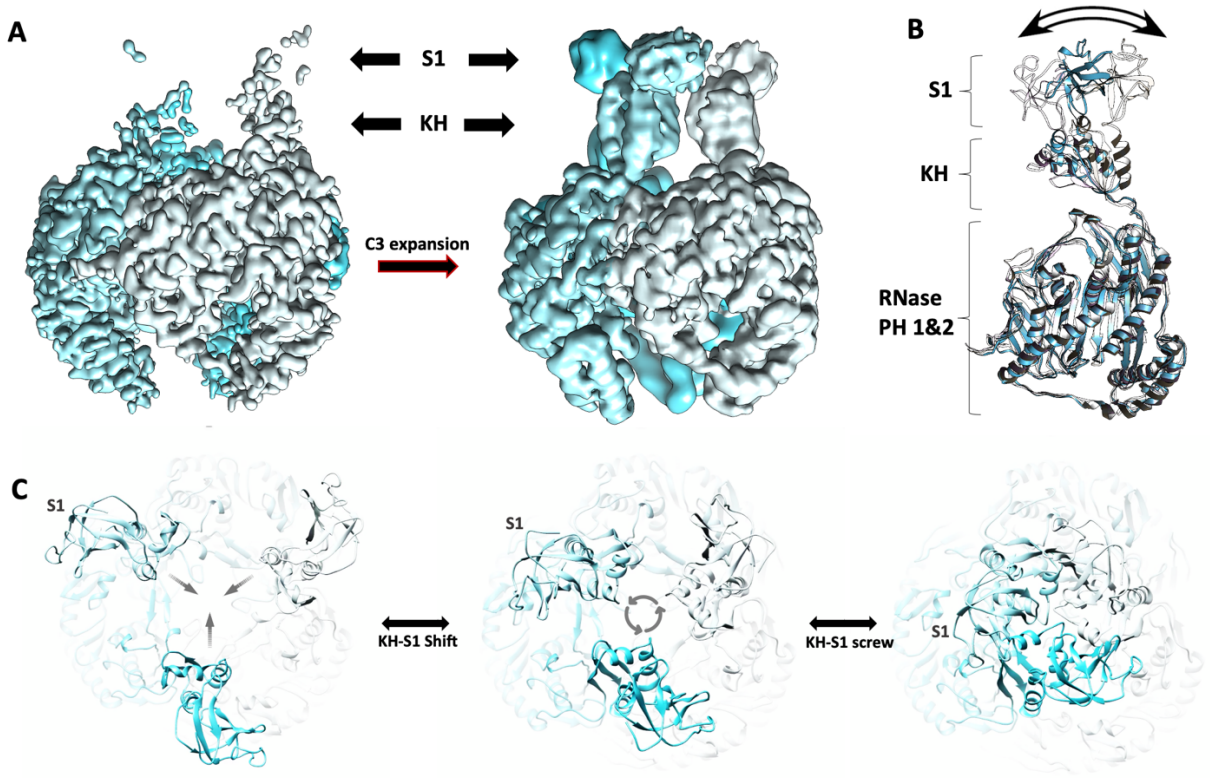


Figure 7: ***Cryo-EM analysis of the apo-PNPase structure***. ***(A)*** *Conventional refinement methods result in poorly resolved densities for the flexible KH-S1 portal. Symmetry expansion (C3) and subsequent masked 3D classification in Relion improve the KH and S1 domain densities.* ***(B)*** *Rigid body docking and molecular dynamics driven fitting of a PNPase homology model in the symmetry expanded map reveals three significantly different states for the KH-S1 modules.* ***(C)*** *If synchronised, these sub-states correspond to opening and closing of the KH-S1 portal via a corkscrew-like motion.*

3.4 The path of 3'ETS^{leuZ} engaged by *E. coli* PNPase in phosphorolytic mode

In the previous section (section 3.3) the apo structure of PNPase was studied and sub-states of the KH-S1 portal resolved. The next step was to study PNPase when engaged on substrate RNAs. Using cryo-EM, I investigated complexes formed between PNPase and several sRNAs under conditions that do not support catalysis (i.e. no phosphate was present in the buffers). RyhB, a sRNA which regulates oxidative stress responses and enzymes of the tricarboxylic acid cycle, and 3'ETS^{leuZ}, which is derived from the 3'UTR of the tRNA operon (Lalaouna *et al.*, 2015) are two of the sRNAs studied in this and the following sections. These two RNAs form a regulatory pair in that 3'ETS^{leuZ} acts as a sponge that sequesters RyhB from its regulatory roles. A third RNA that was studied in complex with PNPase was the sRNA GcvB, which mediates control of numerous transporters (Jin *et al.*, 2009; Urbanowski *et al.*, 2000). Purified GcvB was kindly provided by Dr. Heather Bruce of the Luisi lab. PNPase and the sRNAs were mixed in a 1:1 molar ratio, and 8mM CHAPSO was added to an 8 µM PNPase-sRNA sample (final concentration). Data processing was performed in Relion 3.0.8. Maps for all three PNPase-RNA complexes were reconstructed at sub-4Å resolution (GS-FSC). At this resolution, the three maps were identical to each other and only the results for the PNPase-3'ETS^{leuZ} assembly will be discussed here, for which the global resolution was measured at 3.4 Å (GS-FSC) (Figure 8E, Table 1). Local resolution estimation in Relion shows that the core is at 3.1 Å, but the KH-S1 portal is estimated at 7 Å due to flexibility (data not shown).

Comparison of the maps and models of PNPase in complex with 3'ETS^{leuZ} shows that RNA-binding is associated with compaction of the KH and S1 domains as they clamp on the RNA (Figure 8A). Strikingly, this compaction of the KH-S1 portal is similar to the KH and S1 domain movements found for the apo-PNPase structure, as displayed in Figure 7C. Despite the relatively low resolution of the S1 domains due to conformational heterogeneity, these maps reveal for the first time that the S1 domains come together to cooperatively engage RNA substrates when PNPase is in phosphorolytic mode. The path of the RNA over the KH domains is in agreement with the interactions seen in a co-crystal structure of *C. crescentus* PNPase with RNA (Hardwick *et al.*, 2012). In particular, one out of three KH domains seems to be the main contact point of the RNA on its way to the catalytic core (Figure 7A, C and D). Even though the local resolution near the KH domains doesn't allow for confident modelling of amino acid side chains, there is clear density for the single stranded RNA backbone (Figure 8C) and several potential polar contacts can form between KH residues and the RNA backbone (Figure 8D). For example, Lys571 is part of a GKGG loop in the KH domain and is likely to form hydrogen bonds with the phosphate backbone. This loop and its role in RNA coordination appears equivalent to the GSGG loop in the KH domain of *C. crescentus* PNPase (Hardwick *et al.*, 2012).

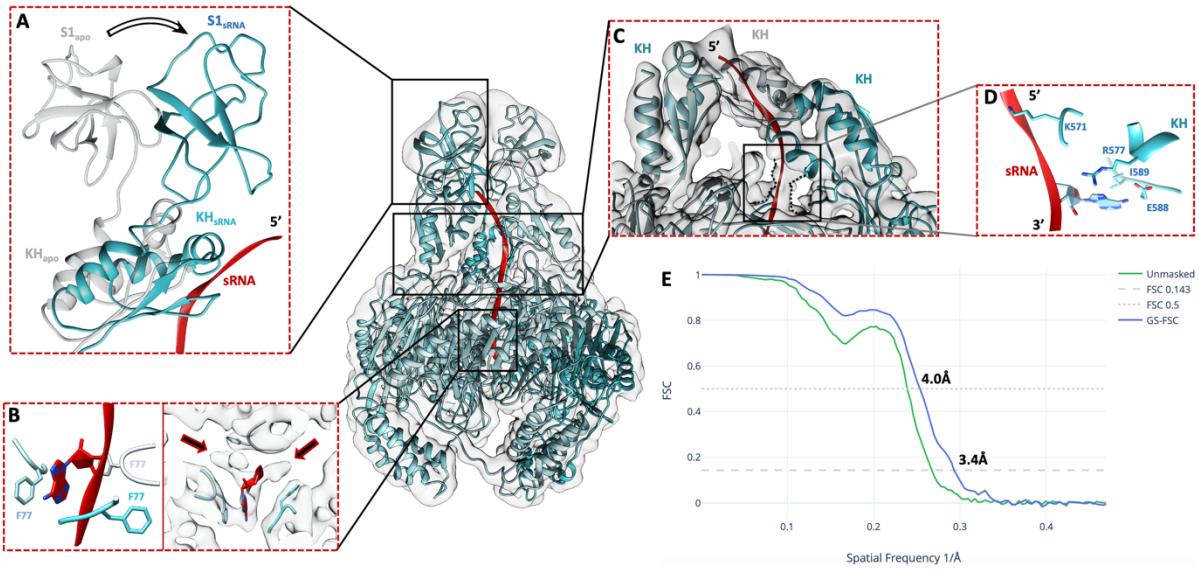


Figure 8: ***CryoEM analysis of PNPase engaged on 3'ETS^{leuZ} substrate. (A)*** *The KH and S1 domains undergo significant repositioning upon RNA binding, wrapping themselves around the RNA substrate in a more compact conformation.* ***(B)*** *Stacking of a base against one of three Phe77 residues marks the entry point into the catalytic PNPase core (left). Extra density was observed for stacking bases against the other Phe77 residues, albeit sterically impossible (red arrows).* ***(C)*** *Single stranded RNA threads towards the core via the KH domains. One out of three KH domains appears to be the main contact point for RNA coordination.* ***(D)*** *Potential polar contacts between residues on the KH domains and the sRNA at the centre of the PNP entry-channel. Lys571 is part of the GKGG loop on the KH domain, analogous to the GSGG loop in C. crescentus PNPase.* ***(E)*** *Global resolution estimation from two independent half maps (GS-FSC). To improve clarity of presentation, the cryo-EM map was lowpass filtered to 7 Å and blurred by 50 Å^2.*

All three maps reveal the same path of the RNA as a single stranded region via the KH-S1 portal and on through the pore entrance to the central channel (Figure 8). The absence of strong density for the RNA 3'-end in the catalytic core could be explained by residual PNPase activity during the cryo-EM sample preparation. Alternatively, the

absence of RNA density in the core could indicate that the 'RNA capture' event does not influence which of three active sites performs the 'RNA cleavage' event, and that any one of three active sites is occupied in each particle, resulting in poorly averaged local densities in the consensus maps. Further support for this hypothesis is apparent from the EM density at the entry channel of the core (Figure 8B). Here, high resolution density for three bases stacking on three Phe77 residues in the entrance channel can be found, which marks the point where the RNA path branches among three possible routes to the active sites (Figure 8B, red arrows). It is sterically impossible for the phosphate backbone to accommodate the stacking of the three bases simultaneously, and it is therefore likely that only one or two bases will stack on the Phe77 residues for any RNA threading into an active site. Interestingly, electron density for three bases stacking against three phenylalanine residues was also observed by Hardwick *et al.* (2012) in their *C. crescentus* PNPase crystal structures. In the same study Hardwick *et al.* propose a threading mechanism for the RNA substrate whereby the KH domains perform rotary movements to direct the RNA to the PNPase core (Hardwick *et al.*, 2012). Possibly, the strain energy from phosphorolysis may be coupled with strand translocation, in analogy to the proposed coupling of RNA hydrolysis and unwinding by the Rrp44 exonuclease from the exosome (Lee *et al.*, 2012). The cryo-EM maps presented here neither confirm nor disprove this hypothesis.

3.5 The PNPase-Hfq-3'ETS^{leuZ} ternary complex reveals a structural basis for the RNA carrier mode

In the previous sections (sections 3.3 and 3.4) cryo-EM models were presented for apo-PNPase and PNPase-sRNA complexes. These structures elucidate how the PNPase KH-S1 portal captures the 3'ETS^{leuZ} substrate. Cryo-EM studies were carried out to obtain a structure of the same PNPase-3'ETS^{leuZ} complex in the protective mode via the addition of Hfq (Bandyra *et al.*, 2016). The sRNA sponge 3'ETS^{leuZ} was observed in *in vivo* pulldown assays of PNPase under conditions of high RyhB expression (Figure 9, carried out by the De Lay group, Department of Microbiology & Molecular Genetics, University of Texas Health Science Center). In addition to the *in vivo* pull-down experiments, gel shift assays demonstrate how a stable and uniform PNPase-Hfq-3'ETS^{leuZ} complex can from *in vitro* (Figure 16B, marked by a red dot).

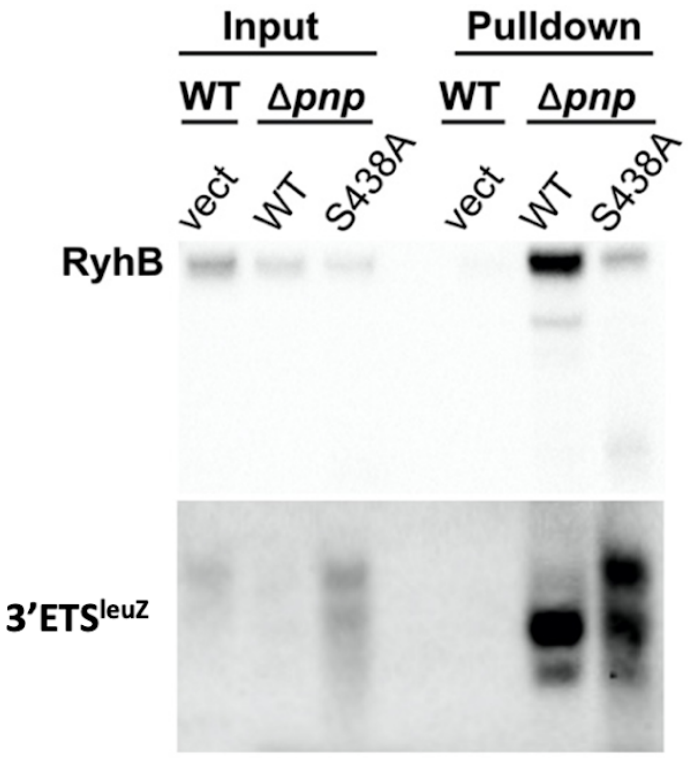


Figure 9: ***PNPase pulldown assays after dipyridyl treatment****, showing that both RyhB (top panel) and 3'ETS^{leuZ} (lower panel) are bound to PNPase after the Pulldown. WT indicates wild type PNPase background and S438A is and inactive PNPase mutant.*

The components of the stable ternary complex of PNPase-Hfq-3'ETS^{leuZ} were mixed in the absence of phosphate to a final concentration of 10 µM for the assembly. 8 mM CHAPSO was added to the sample right before grid freezing, resulting in single particles without any orientation bias (Figure 10A). 2D classification resulted in classes with a well-defined catalytic core for PNPase and a distinct region of high conformational heterogeneity at the KH-S1 portal (Figure 10B). Consensus refinements revealed that the alignments are dominated by the rigid and symmetric PNPase core, which results in a well resolved density for the core domains (3.2 Å) and diffuse, uninterpretable density for the KH-S1 portal region (Figure 10C).

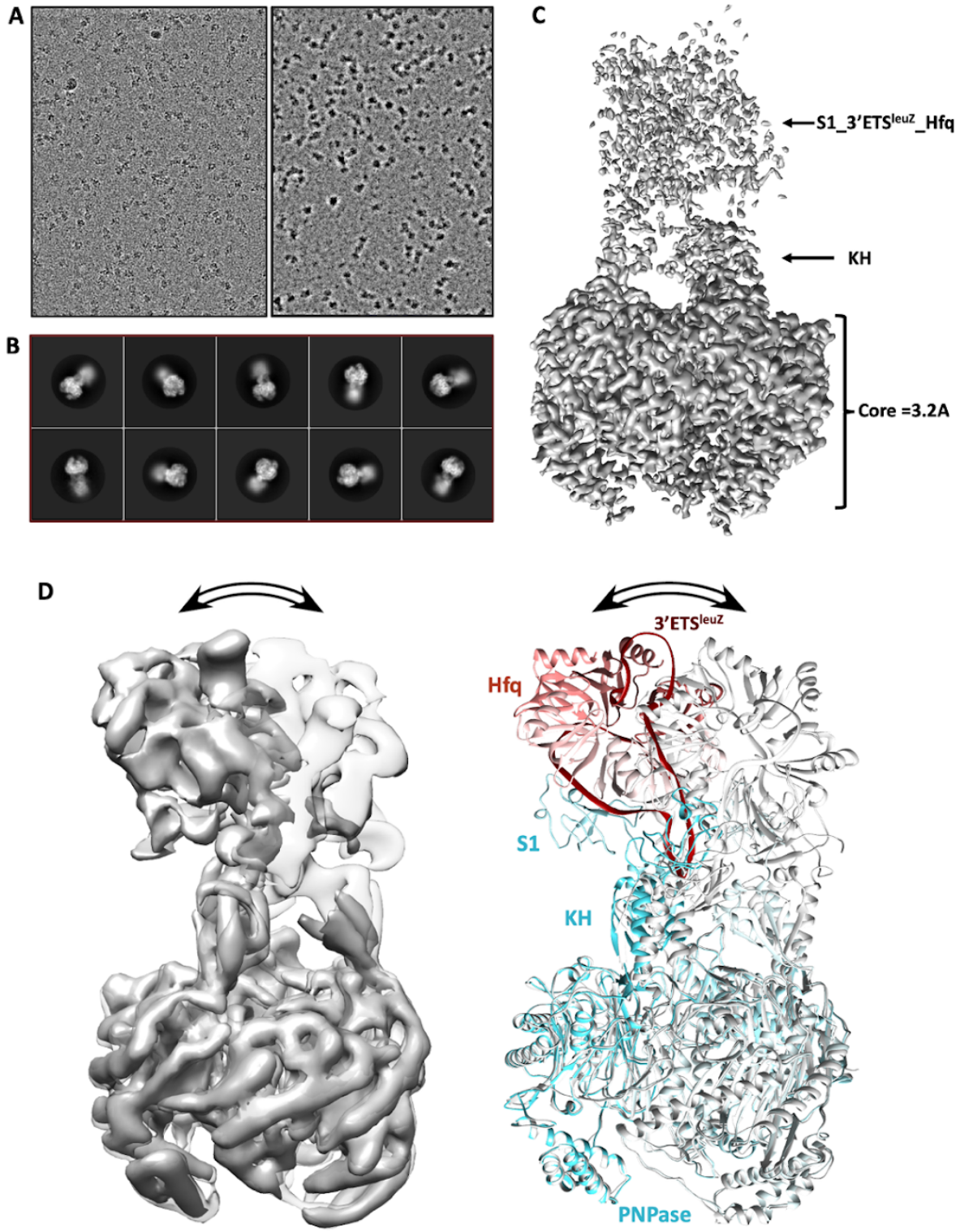


Figure 10: ***CryoEM analysis of the PNPase-Hfq-3'ETS^{leuZ} complex***. ***(A)*** *Representative micrographs collected without the Volta phase plate (VPP) (left, 3 µM defocus) and with VPP) (right 0.5 µM defocus).* ***(B)*** *2D class averages show a well resolved core and a diffuse KH-S1/Hfq portal.* ***(C)*** *Consensus 3D refinements reveal a well resolved core and diffuse densities at the KH-S1 portal.* ***(D)*** *The PNPase-Hfq-3'ETS^{leuZ} assembly can adopt multiple conformations, as the Hfq-3'ETS^{leuZ} density can occupy different states along a continuous clapper-like trajectory. Due to significant internal movements of the KH and S1 domains, and the small size of Hfq (67 kDa), a rigid body approach did not yield significant increases in local resolution. Instead, careful 3D classification approaches in Relion and cryoSPARC allowed for the separation of particle populations belonging to distinct states along the*

continuous trajectory, improving the interpretability significantly. Three conformational states could be resolved, of which the two extremes are depicted here. Maps are lowpass filtered to 7 Å for to improve visibility.

Extensive 3D classification in Relion and cryoSPARC revealed a series of conformational sub-states and well resolved density for one Hfq chaperone near the KH and S1 domains (Figure 10D, Material and Methods section 5.7). A detailed overview of the PNPase Hfq-3'ETS^{leuZ} data processing pipeline is depicted in Figure 11. The density for the KH and S1 domains as well as the Hfq and sRNA in each conformational sub-class improved significantly after combining four datasets, two of which were collected with a Volta Phase Plate (VPP) (Figure 11 and Figure 10A, Table 1) (Danev & Baumeister, 2016). Focused 3D classification and/or masked 3D refinement approaches in Relion did not further improve the maps due to internal movements of the KH-S1 domains and the small size of Hfq (67 kDa). In parallel, 3D variability analyses were carried out in cryoSPARC to further explore the conformational landscape of the PNPase-Hfq-3'ETS^{leuZ} RNA carrier assembly (Figure 11 and Figure 12C). During 3DVA the resolution was limited to 6 Å, i.e. only structural variability up to 6 Å was considered, and six modes of variability were solved. Two modes of molecular motion were readily interpretable, corresponding to a 'rocking' and 'rotary' motion of the bound Hfq chaperone (Figure 12C). Notably, the molecular motions revealed by 3DVA are in good agreement with the discrete classes resolved by 3D classification approaches in Relion and cryoSPARC (Figure 10D, Figure 12C).

The best PNPase-Hfq-3'ETS^{leuZ} class was refined in cryoSPARC to an overall resolution of 3.7 Å, locally ranging from 3.2 Å in the PNPase core to 4.4 Å - 7.2 Å in the flexible KH-S1 portal (Figure 13A and B, Figure 11, state 1, Table 1). At these resolutions, the path for the 3'ETS^{leuZ} RNA backbone was clearly defined, with interpretable density for the bases in the best regions (Figure 13C). Secondary structures were well resolved for all three KH domains and one S1 domain. Docking and refinement of models in the different conformational subclasses reveal likely modes of conformational switching (Figure 10D). Although the movements may not be mechanistically relevant for the equilibrium complex, they could play a role in exposing the KH-S1-Hfq portal for potential handover of the 3'ETS^{leuZ} tRNA operon fragment to its target RNAs.

.

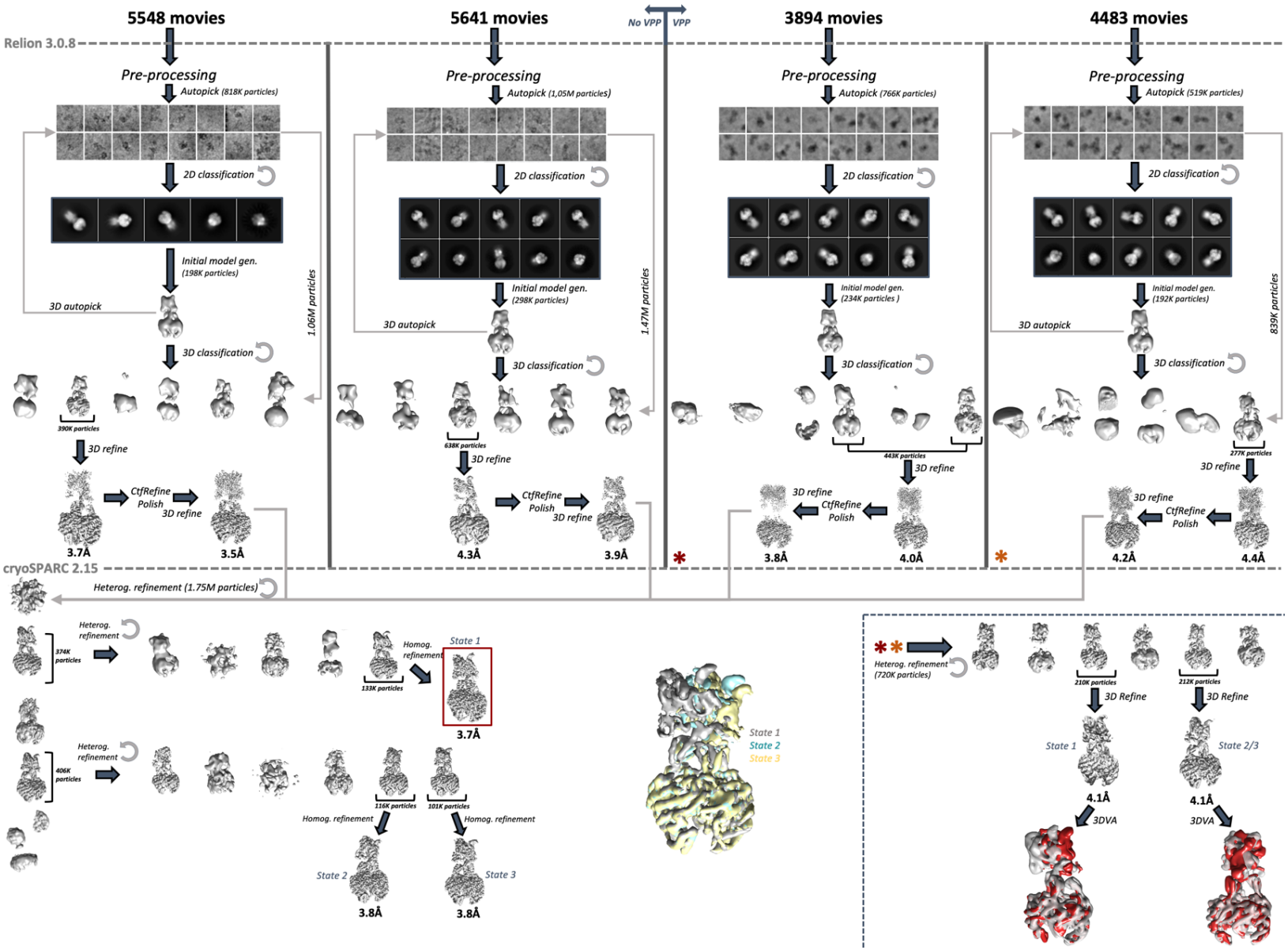

5548 movies
5641 movies
No VPP
VPP
3894 movies
4483 movies
Relion 3.0.8
Pre-processing
Autopick (818K particles)
Autopick (1,05M particles)
Autopick (766K particles)
Autopick (519K particles)
2D classification
Initial model gen. (198K particles)
Initial model gen. (298K particles)
Initial model gen. (234K particles)
Initial model gen. (192K particles)
3D autopick
1.06M particles
1.47M particles
839K particles
3D classification
390K particles
638K particles
443K particles
277K particles
3D refine
CtfRefine
Polish
3.7Å
3.5Å
4.3Å
3.9Å
3.8Å
4.0Å
4.2Å
4.4Å
cryoSPARC 2.15
Heterog. refinement (1.75M particles)
Heterog. refinement
374K particles
406K particles
133K particles
116K particles
101K particles
Homog. refinement
State 1
State 2
State 3
3.7Å
3.8Å
3.8Å
Heterog. refinement (720K particles)
210K particles
212K particles
3D Refine
State 2/3
4.1Å
3DVA

Figure 11: ***Schematic overview of the processing pipeline for PNPase-Hfq-3'ETS^{leuZ}.*** *For all four datasets pre-processing, which entails motion correction, ctf estimation, manual picking and 2D reference generation was carried out in Relion 3.0.8. 2D/3D classifications were carried out in Relion too, as well as consensus 3D refinements and per particle ctf refinement/particle polishing. Clean particle sets for each dataset we transferred to cryoSPARC and combined. Subsequent rounds of heterogenous refinements were used to resolve three different conformational states (state 1, state 2 and state 3), each of which refined to sub-4 Å resolution. In parallel, conformational heterogeneity was also resolved through heterogeneous refinement for the particle sets from the VPP-datasets (see *, bottom right inset). To further analyse the conformational landscape for each state, 3D variability analysis was carried out on state1 and state 2/3, resolving two dominant modes of molecular motion.*

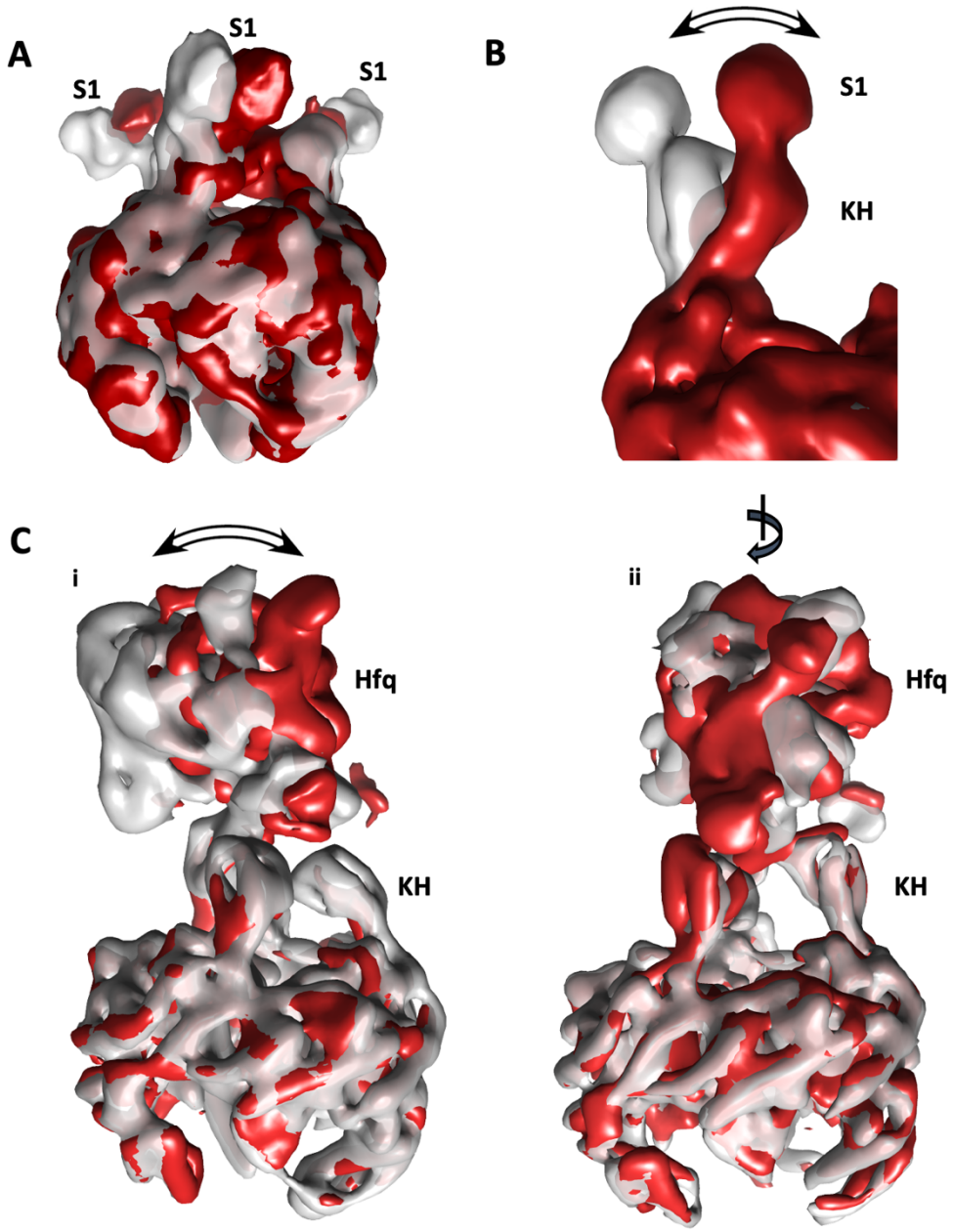


Figure 12: ***3D Variability Analysis of the apo-PNPase and PNPase-Hfq-3'ETS^{leuZ} particles. (A)*** *cryoSPARC 3D Variability Analysis (3DVA) of apo-PNPase after C3 symmetry expansion (Punjani et al., 2017; Punjani & Fleet, 2020). Grey and red maps represent extreme states of the molecular motion of the KH and S1 domains as encoded in the dominant 3DVA mode.* ***(B)*** *Zoom-in on one KH-S1 region, showing conformational heterogeneity of these domains as encoded in the dominant 3DVA mode.* ***(D)*** *PNPase-Hfq-3'ETS^{leuZ} particles were subjected to 3DVA in cryoSPARC to investigate the conformational landscape. This approach revealed a 'rocking' mode (i) and a 'rotary' mode (ii) of conformational variability in the Hfq-3'ETS^{leuZ} region. All maps presented here are lowpass filtered to 8 Å for visibility.*

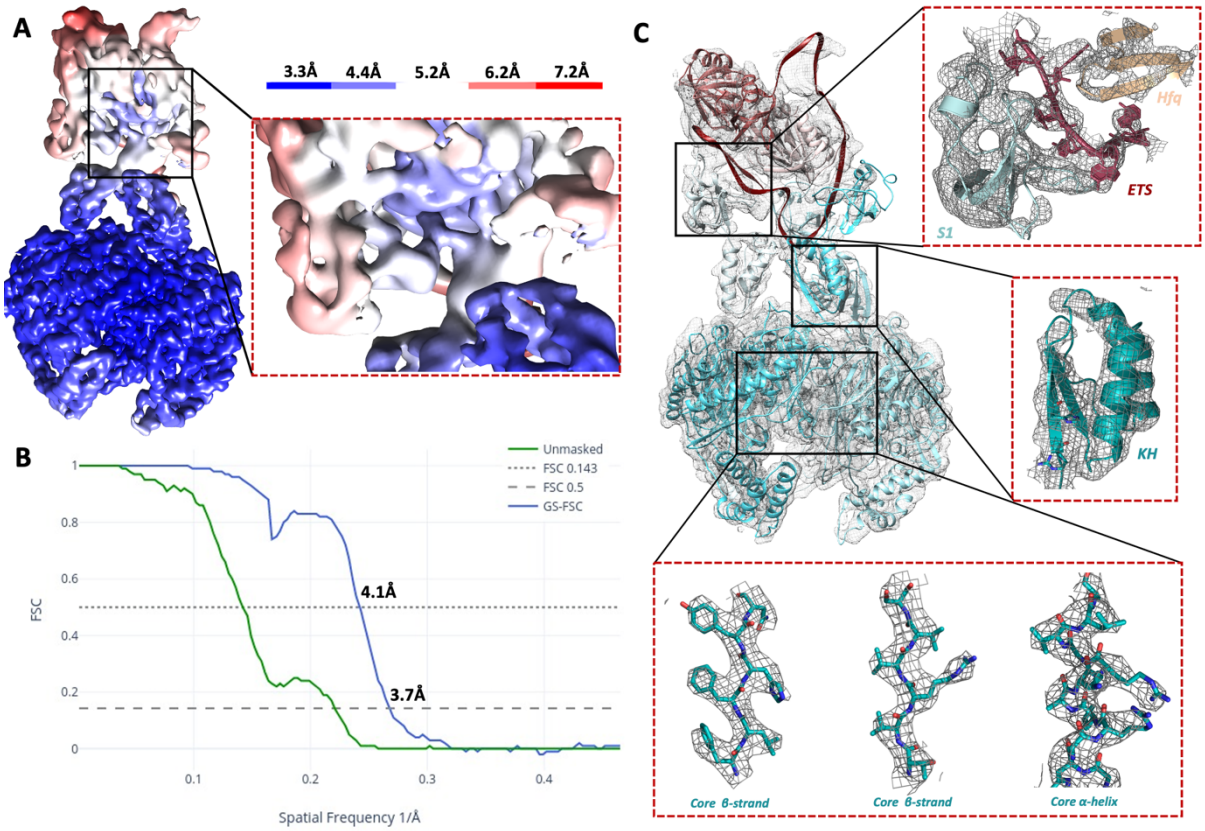


Figure 13: ***Global and Local resolution estimation for the PNPase-Hfq-3'ETS^{leuZ} map***. ***(A)*** *Local resolution estimation and filtering as performed in cryoSPARC. The PNPase core is rigid and was reconstructed at 3.2Å. The KH, S1, 3'ETS^{leuZ} and Hfq densities suffer from a loss in resolution due to severe flexibility, with local resolutions ranging from 4.4 to 7.2Å after careful 3D classification. Local resolutions were calculated from reconstructed half maps, at FSC$_{0.5}$.* ***(B)*** *Global resolution estimation from independent half-maps (GS-FSC).* ***(C)*** *Selected fits of the model in the cryo-EM density at the PNPase core and KH-S1/Hfq portal.*

The overall architecture of the assembly can be readily observed in the cryo-EM models. The 3'ETS^{leuZ} RNA in the ternary complex is engaged on the distal and proximal surfaces of Hfq, but also interacts with the circumferential rim (Figure 14A and B). Although the resolution is limited in this region it is clear that the PNPase KH-S1 portal coordinates the sRNA extensively and the individual KH and S1 domains form many putative contacts with 3'ETS^{leuZ}. A short, exposed hairpin on the Hfq distal side threads into the KH-S1 portal. Here the hairpin is mainly bound by a basic loop extending from two PNPase S1 domains (Figure 14B and C, S1_3/S1-loop3 and S1_2/S1_loop2). The same hairpin is in close proximity to basic residues on all three KH domains (Figure 14B and C, KH-1, KH-2 and KH-3), where the RNA folds back, away from the active site. The S1-2 and S1-3 domains also coordinate the 5' end of 3'ETS^{leuZ} (Figure 14B and F) and a single stranded region of the RNA goes over the Hfq rim (Figure 14B and E), respectively. A third S1 domain (S1-1) extensively coordinates an A-rich region of the 3'ETS^{leuZ} RNA bound to the Hfq distal side (Figure 14B and D, S1-1 and S1-loop1), and is the best resolved S1 domain in the cryo-EM map (Figure 13C). Overall the same residues on the three KH domains are involved in RNA binding, which is also the case for the individual S1 domains (see next paragraph). Interestingly, the KH and S1 domains adopt an open conformation that resembles more closely the conformation observed in the apo-state than in the substrate-bound mode described in the previous sections (Figure 7C and Figure 8).

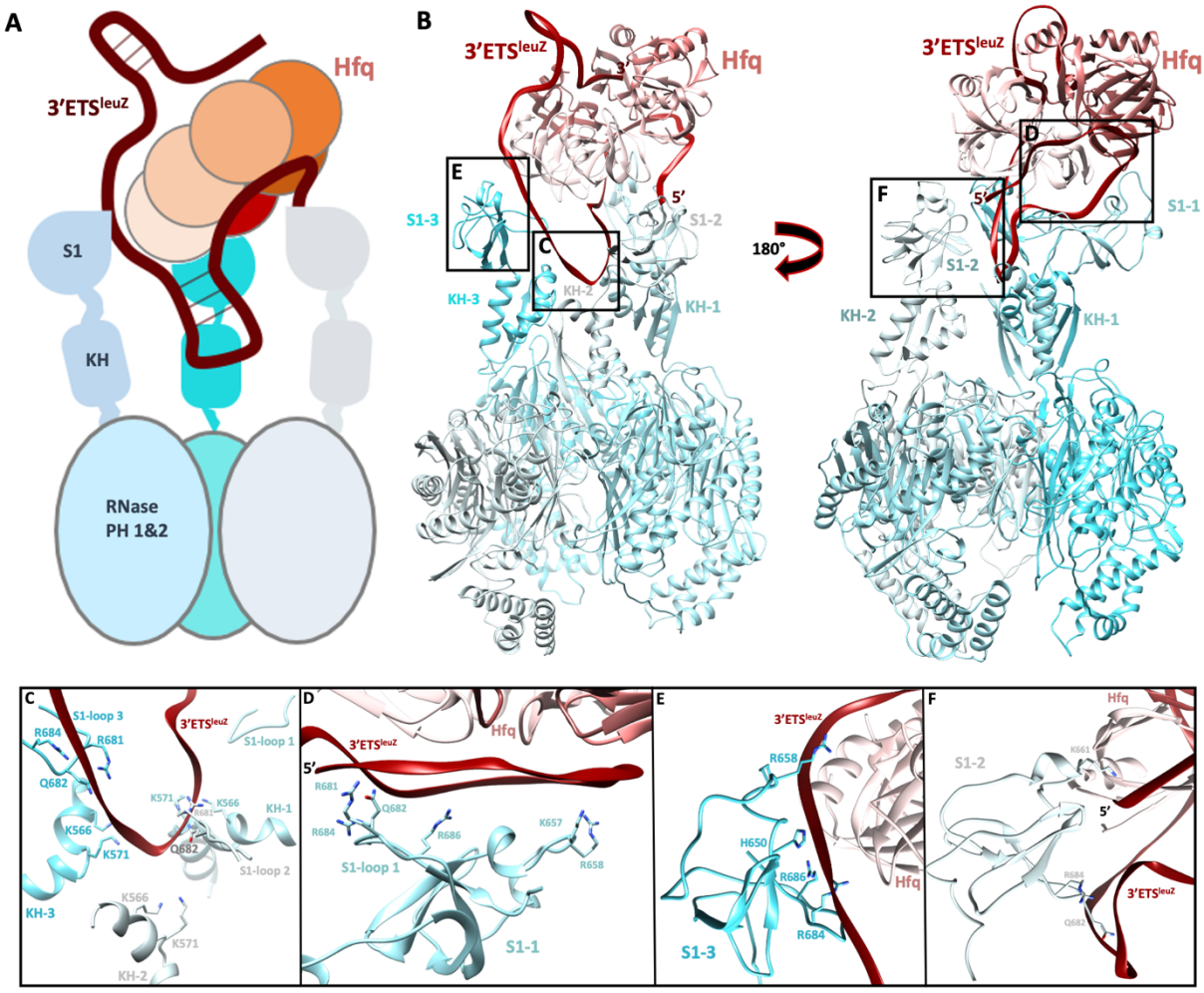


Figure 14: ***Architectural principles of the PNPase RNA carrier mode***. ***(A)*** *Schematic model of the PNPase-Hfq-3'ETS^{leuZ} complex.* ***(B)*** *Structure of the PNPase-3'ETS^{leuZ}-Hfq RNA carrier assembly. Rerouting of the sRNA by Hfq prevents it from being degraded by PNPase. The KH and S1 domains are crucial for complex formation on Hfq-sRNA. No direct protein-protein contacts between PNPase and Hfq were found.* ***(C)*** *All three KH domains and a basic loop on two S1 domains (S1-loop 2 and S1-loop 3) coordinate a 3'ETS^{leuZ} stem-loop on the Hfq distal side.* ***(D)*** *The same basic loop on a third S1 domain (S1-1), containing R681, Q682, R684 and R686, coordinates the 3'ETS^{leuZ} RNA bound to the Hfq distal side, in addition to S1 residues K657 and R658.* ***(E)*** *S1-3 helps coordinate the sRNA as it threads over the Hfq rim towards the proximal side.* ***(F)*** *S1-2 appears to bind and coordinate the 5'-end of 3'ETS^{leuZ} in the cryo-EM model.*

Table 1: Cryo-EM data collection and refinement statistics for PNPase structures

Structure	PNPase	PNPase-ETS	PNPase-Hfq-3'ETS^{leuZ}
Data collection			
Microscope	FEI Titan Krios	FEI Titan Krios	FEI Titan Krios
Voltage (kV)	300	300	300
Detector	Gatan K2	Gatan K2	Gatan K2
Nominal magnification	130 000	130 000	130 000
Pixel size (Å)	1.065	1.065	1.065
Electron dose, per frame (e^{-}/Å^2)	1.40	1.41	1.45

Electron dose, total (e^{-}/Å^2)	53.0	53.6	54.1
Defocus range (μm)	-1 / -2.5	-1 / -2.5	-1 / -2.5
Exposure (s)	12	12	12
Frames	38	38	38
Number of micrographs	1008	2741	19566
Reconstruction			
Software	RELION-3.0.8	RELION-3.0.8/cryoSPARC 2.15	RELION-3.0.8
Number of particles used	58 000	206 803	133607
Final resolution, $FSC_{0.143}$ (Å)	3.4	3.4	3.7
Map-sharpening B factor (Å^2)	-76	-110	-123
Model composition			
Non-hydrogen atoms	15921	16091	19918
Protein residues	2085	2085	2454
RNA nucleotide	0	8	50
Molar Mass (kDa)	231	253	318
Refinement			
Software	Refmac5/Phenix/Isolde	Refmac5/Phenix/Isolde	Refmac5/Phenix/Isolde
Correlation coefficient, masked	0.85	0.84	0.84
Correlation coefficient, box	0.88	0.86	0.86
$FSC_{0.5}$ (model-map)	3.4	3.6	3.8
Validation			
MolProbity score	1.44	1.33	1.7
Clash score, all atoms	3.16	2.23	6.4
Ramachandran statistics			
Favoured, overall (%)	95.27	95.24	95.73
Allowed, overall (%)	4.49	4.62	3.98
Outlier, overall (%)	0.24	0.14	0.29
R.m.s. deviations			
Bond length (Å)	0.0064	0.014	0.0072
Bond angle (°)	1.55	1.55	1.52

3.5.1 *In vitro* and *in vivo* validation of RNA carrier complexes

Although the resolution in the KH-S1/Hfq region in the PNPase-Hfq-3'ETS^{leuZ} assembly is limited due to conformational heterogeneity, the model allowed to predict and test candidate residues in the PNPase domains that contact the sRNA in the RNA carrier complex. To test the importance of these interactions, three PNPase mutants were designed and evaluated for RNA and RNA-Hfq binding *in vitro*: (1) A double mutant of residues

K657A and R658A in the S1 domain involved in the interaction with 3'ETS^{leuZ} on the distal side of Hfq (S1 domain; PNPase S1x2), (2) a double mutant of the KH domain residues K566A and K571A involved in binding the 3'ETS^{leuZ} stem loop (KH domain; PNPase KHx2), and (3) a quadruple mutant of S1 domain residues R681A, Q682A, R684A, R686A involved in the interaction with 3'ETS^{leuZ} (S1 domain; PNPase S1x4) (Figure 14C-F). All the mutant proteins were expressed using a Δ*pnp* strain to avoid formation of heterotrimers with wild type subunits from the host. Three sRNAs were chosen to study the effects of these mutations: the 3'ETS^{leuZ} captured in the cryo-EM structures described above, and two additional sRNAs representing the two classes of Hfq-binding sRNAs, namely RyhB (class I) and CyaR (class II), which have been shown to form RNA carrier complexes with PNPase (De Lay & Gottesman, 2009; Johansen *et al.* 2008; Bandyra *et al.*, 2016). The activity assays and gel-shift assays presented next were carried out in close collaboration with Dr. Kasia Bandyra, Luisi group, Department of Biochemistry, University of Cambridge.

All purified PNPase mutants were active, albeit to a lesser extent than wild type enzyme, in agreement with the hypothesis that the KH and S1 domains support substrate capture, but are not directly involved in phosphorolytic RNA digestion (Figure 15A-C, panels i). In the presence of Hfq, none of the PNPase mutants nor the wild type enzyme can degrade the sRNAs, suggesting Hfq masks the RNA from the PNPase catalytic core (Figure 15A-C, panels ii). Notably, efficient PNPase-Hfq-sRNA ternary complex formation was not observed in subsequent electrophoretic mobility shift assays (EMSA) for any of the PNPase mutants. The PNPase S1x2 and KHx2 mutants still had reduced affinity for some sRNA-Hfq pairs but the PNPase S1x4 mutant did not form ternary complexes with Hfq and any of the tested sRNAs (Figure 16A-C). These results are in agreement with the proposed model for the RNA carrier mode of PNPase, which predicts that the basic loop in the S1 domain is paramount in formation of the assembly (Figure 14B-F).

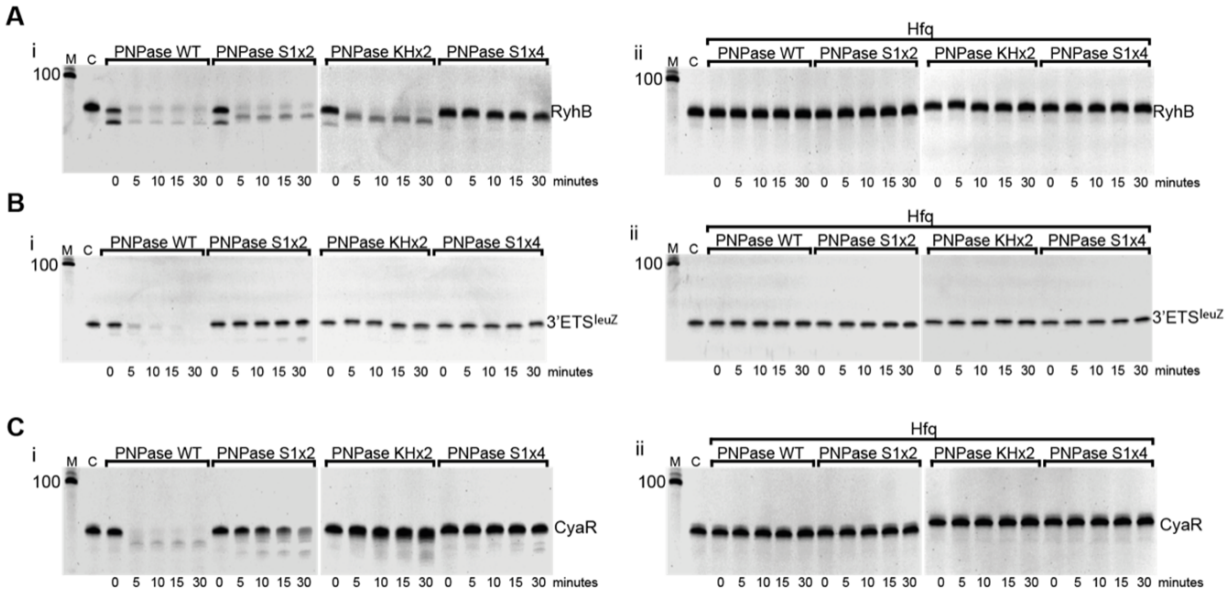


Figure 15: ***PNPase WT and mutants (S1x2, KHx2 and S1x4) display different levels of activity towards sRNAs****. Degradation of RyhB (**A**), 3'ETS^{leuZ} (**B**) and CyaR (**C**) by PNPase in the absence (i) or presence (ii) of Hfq. M is molecular weight marker; C is control RNA without PNPase treatment. Timepoints are depicted at the bottom of each gel.*

To determine the contribution of theKH-S1 domains in the context of the RNA carrier assembly to sRNA stability *in vivo*, the impact of the PNPase mutants S1x2, KHx2 and S1x4 on CyaR and RyhB stability was examined after inhibition of transcription initiation by rifampicin treatment. These experiments were carried out by the De Lay group, Department of Microbiology & Molecular Genetics, University of Texas Health Science Center. As shown in Figure 16D and Table 2, all three sets of substitutions resulted in increased turnover in RyhB, i.e. its half-life decreased from 14 min in the strain expressing PNPase-3xFLAG to 4.1, 4.1, and 6.5 min for those producing PNPase S1x2, KHx2 and S1x4, respectively. CyaR stability significantly decreased in strains expressing PNPase KHx2 and S1x4, whereas expression of PNPase S1x2 caused a modest reduction in stability compared to the PNPase-3xFLAG control (Figure 16E). Next, we compared the ability of the wild type and mutant forms of the PNPase constructs to interact with RyhB and CyaR via PNPase co-immunoprecipitation assays. These experiments show that the PNPase S1x4 mutant was defective in binding CyaR and RyhB, whereas the PNPase S1x2 and KHx2 pulled down similar amounts of these RNAs as the PNPase-3xFLAG (Figure 16F and G). Altogether, these *in vivo* findings are consistent with the *in vitro* results indicating that a positively charged loop within the S1 domain formed by R681, Q682, R684, and R686 is important for sRNA interaction and stabilization *in vivo*, probably due to its indispensability in RNA carrier complex formation, and that positively charged residues in the KH domain have a significant, but less substantial role in the RNA carrier assembly.

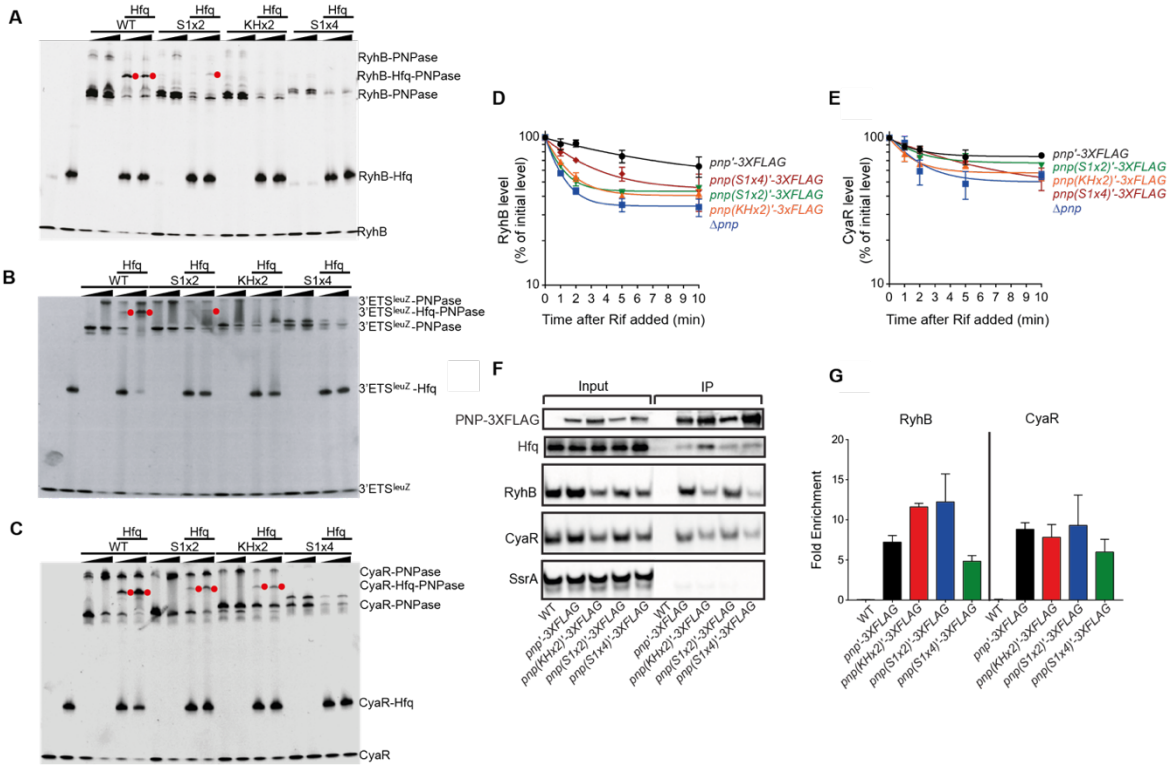


Figure 16: ***The KH-S1 portal is crucial for PNPase-sRNA-Hfq complex formation.*** *EMSA of wild type PNPase and KH-S1 mutants with RyhB* ***(A)****, 3'ETS^{leuZ}* ***(B)****, and CyaR* ***(C)*** *in the absence and presence of Hfq. Ternary complexes are highlighted with a red dot. Two different PNPase concentrations were used for every PNPase construct (1:1 and 1:3 RNA:PNPase molar ratio), represented by gradient bars on top of the gels.* ***(D, E)*** *RNA half-life experiments to determine RyhB and CyaR sRNA stabilities in an E. coli strain expressing a 3X-FLAG tagged construct of PNPase WT and mutants. RyhB and CyaR signal intensities were quantified using northern blots and normalized to their corresponding loading controls (SsrA). sRNA decay curves were generated by fitting*

the normalized signal intensities for each time point. Points and error bars in the curves represent the means and the standard errors (SEM) of at least three independent experiments. RyhB and CyaR half-life measurements corresponding to RNA stability curves are listed in Table 2. ***(F, G)*** *Cell extracts prepared from late exponential phase cultures of E. coli strains expressing WT PNPase, or Flag-tagged PNPase WT and mutants were used to assess coprecipitation of sRNAs, which were analysed by northern blot.* ***(G)*** *Fold enrichment of a given RNA upon immunoprecipitation was determined by first calculating the signal intensity per microgram of RNA for the input and the elution from the northern blots in (**F**). The normalized elution signal was then divided by the input signal. An untagged wild-type strain (WT) was used as a control for data presented in* ***(F, G)****. S1x2 – PNPase K657A, R658A; KHx2 – PNPase K566A, K571A; S1x4 – PNPase R681A, Q682A, R684A, R686A.*

Table 2: sRNA Half-lives from *Figure 16*D and E

sRNA	Strain genotype	Avg. half-life (min) ± SE
CyaR	*pnp'*-3XFLAG	17.5 ± 4.2
	Δ*pnp*	6.8 ± 1.9
	pnp(2x1)'-3xFLAG	12.7 ± 1.9
	pnp(2x2)'-3xFLAG	8.1 ± 1.5
	pnp(4x)'-3xFLAG	9.8 ± 1.3
RyhB	*pnp'*-3XFLAG	14.0 ± 2.7
	Δ*pnp*	2.5 ± 0.72
	pnp(2x1)'-3xFLAG	4.1 ± 1.1
	pnp(2x2)'-3xFLAG	4.1 ± 0.90
	pnp(4x)'-3xFLAG	6.5 ± 1.1

3.5.2 An imperfect ARN-repeat sequence supports the Hfq-PNPase interface

In the 3'ETS^{leuZ} RNA carrier assembly, no direct protein-protein contacts were found between PNPase and Hfq. Instead, 3'ETS^{leuZ} mediates the association of the two proteins. Strikingly, the interactions between 3'ETS^{leuZ} and Hfq resemble closely the proposed binding mode of type II sRNAs, where the RNA is predicted to bind the distal surface via an adenine-rich stretch in its sequence and the proximal side via its U-rich 3' tail (Schu *et al.*, 2015; Zhang *et al.*, 2013). Indeed, the 3' end of the 3'ETS^{leuZ} RNA, marked by a poly U sequence, is bound to the proximal face, and is directed away from the path that would lead to the central channel in the PNPase degradative mode (Figure 17A). The 3'ETS^{leuZ} backbone could be traced over the rim of the Hfq chaperone, where extra cryo-EM density, albeit weak, was resolved for a C-terminal tail of one of the Hfq protomers (Figure 17A and B). Even though the backbone of only a few residues of the Hfq C-terminus could be traced in the map, these observations suggest that the flexible, C-terminal Hfq tail binds substrate RNAs on the Hfq rim and proximal side.

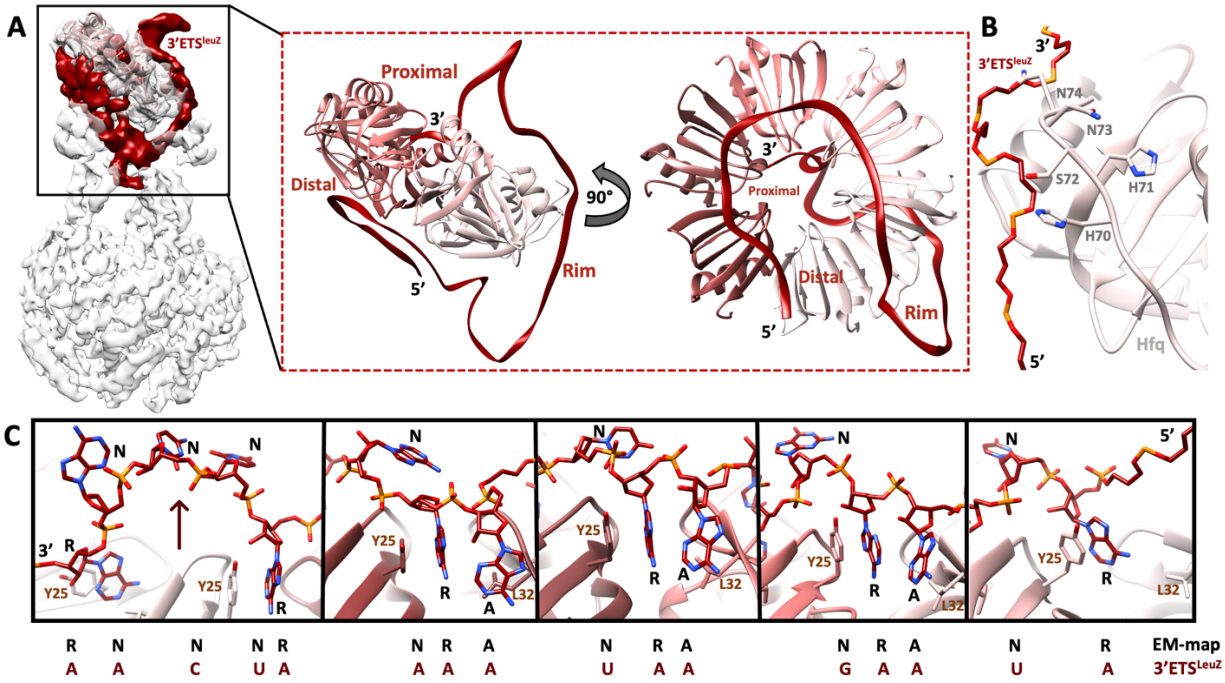


Figure 17: ***3'ETS^{leuZ} interactions with Hfq (A)*** *3'ETS^{leuZ} binds the Hfq distal side, proximal side and rim. The parts of 3'ETS^{leuZ} on the distal side and rim are presented to PNPase.* ***(B)*** *Interpretable density was observed for part of the C-terminal tail of one of the Hfq protomers. Polar (Ser72, Asn73-74) and charged (His 70-71) residues on the Hfq C-terminal tail could help coordinate the RNA on the Hfq rim-proximal side cross-over.* ***(C)*** *An incomplete and degenerate ARN-repeat motif is bound as a ring-like "Brennan crown" fold on the Hfq distal side, as observed previously for a perfect 6xARN motif (Link et al. 2009; Pei et al., 2019). All R pockets are occupied yet only three A pockets are used. Regions where the corresponding base doesn't fit the A pocket, the 3'ETS^{leuZ} backbone is slightly detached from the Hfq distal side and interacts with PNPase S1 domains (left panel, dark red arrow shows backbone 'lift off'). The map in panel A is locally filtered according to the estimated local resolutions and B-factor blurred for presentation.*

The 3'ETS^{leuZ} RNA binds the Hfq distal face with the same interaction motif observed for other Hfq complexes in gram-negative bacteria (Pei *et al.*, 2019, Link *et al.*, 2009). This motif, thematically named the "Brennan crown", has been described extensively in the previous chapter (Chapter I, Architectural principles of Hfq-Crc mediated gene regulation), where it was found to mediate the interaction between Hfq and Crc in *Pseudomonads* to regulate translation of target genes. The motif consists of an A-R-N repeat pattern (A is adenine, R is a purine, N is any nucleotide) and has A and R buried in dedicated pockets on the Hfq distal face, while the N base is presented outward and exposed (Figure 17C). Interestingly, the cryo-EM density for the PNPase-3'ETS^{leuZ}-Hfq assembly suggests the Brennan fold is both incomplete and degenerate for 3'ETS^{leuZ}. In particular, closer inspection of the Hfq region in the cryo-EM map shows that not all six 'A' pockets on the Hfq distal side are occupied, as cryo-EM density could be observed for only three A-site bases. In contrast, all six R pockets are occupied on the Hfq distal side and each of the 3'ETS^{leuZ} R-site purines is coordinated by Tyr25 via base stacking (Figure 17C). Combining these observations, only three complete ARN triplets could be modelled into the density, yielding a degenerate Brennan pattern of 5' RN ARN ARN ARN RNN NR 3' (Figure 17C, Figure 18). These 'deviations' from the ARN-rule observed in the cryo-EM maps allow the 3'ETS^{leuZ} backbone to 'lift off' the Hfq distal face when interacting with one of the PNPase S1 domains (Figure 17C, left panel). In agreement with the cryo-EM model, the 3'ETS^{leuZ} sequence lacks a

complete set of six ARN motifs that would have the potential to fully occupy the distal face of Hfq. Interestingly, the 3'ETS^{leuZ} sequence harbours a single 5' RN ARN ARN ARN RNN NR 3' motif at its 5' end, which is where the motif is expected to be located according to the cryo-EM map, directly adjacent to the RyhB seed region (Figure 17C, Figure 18). Consequently, the seed region is partially wrapped in the KH-S1 portal of PNPase and partially presented on the Hfq rim.

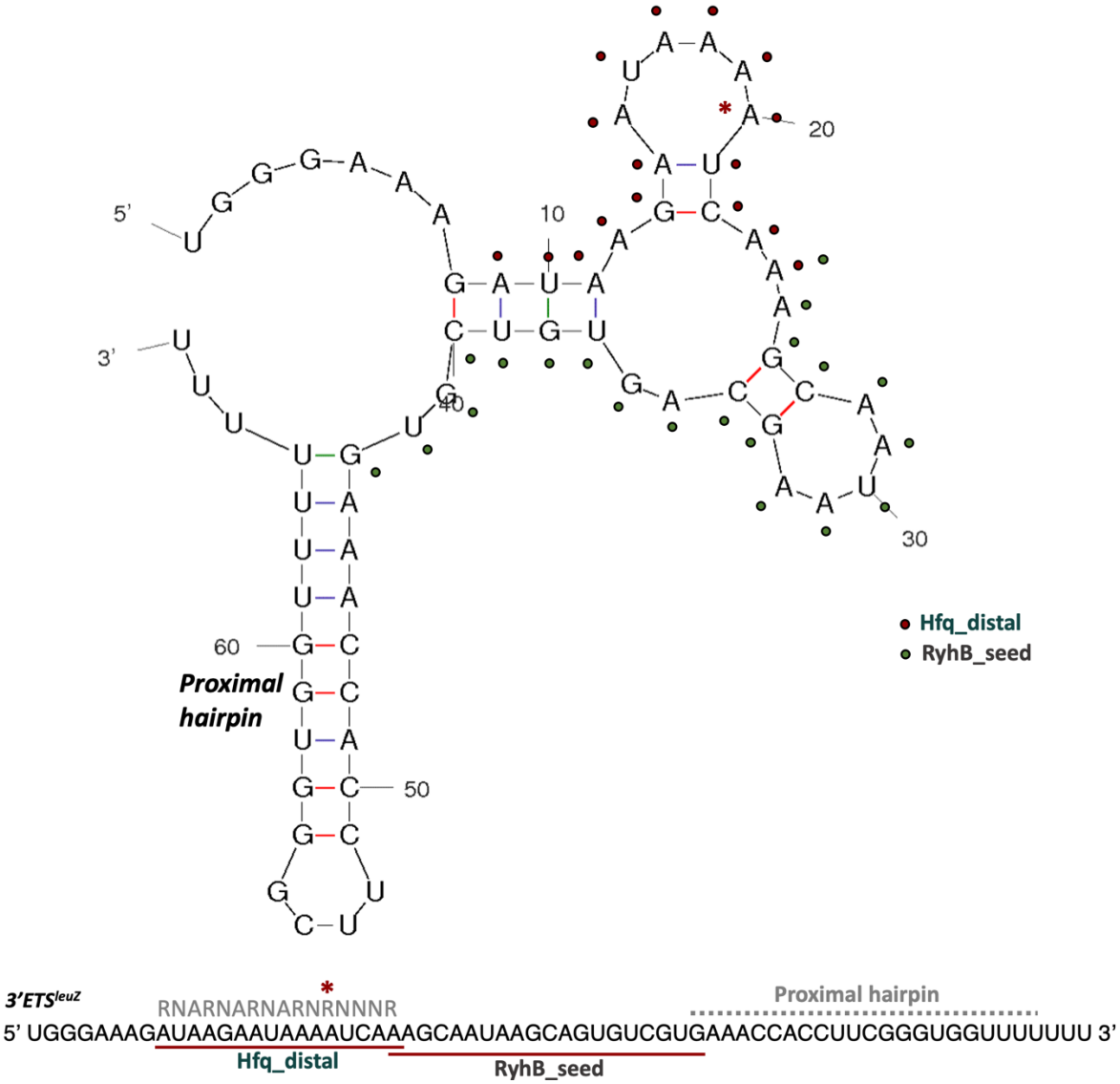


Figure 18: ***Annotated secondary structure prediction of 3'ETS^{leuZ}.*** *Red dots indicate which 3'ETS^{leuZ} residues are presented to the PNPase S1 domains by the Hfq distal side. Green dots mark the RyhB seed sequence. An annotated sequence is depicted at the bottom of the figure. The sequence that was modelled in the cryo-EM map is underlined in red. Occupied A-, R- or N- sites are annotated in grey above the modelled sequence. * refers to an A-site 'skipping'-violation of the ARN-rule, where the A-site pocket is left empty on the Hfq distal side. The range for the hairpin on the Proximal side of Hfq is arbitrary due to low local resolutions in the corresponding map regions. The secondary structure was predicted with mfold (Zuker, 2003).*

The observations described above confirm, for the first time, the fold and interaction pattern of class-II sRNAs when bound to Hfq. In addition, these results show how an incomplete and degenerate 'ARN'-triplet repeat region is still sufficient for highly specific assembly formation between an sRNA presented on the Hfq distal side and

PNPase. This in turn indicates that presentation of substrate RNAs on the Hfq distal side allows rather degenerate ARN-sequence codes, as was observed in Chapter I as well. If true, the range of substrate RNAs that could be regulated by Hfq-mediated effector complexes analogous to the RNA carrier complex between PNPase, 3'ETS^{leuZ} and Hfq is much broader than initially thought.

3.6 Longer Type II sRNAs can accumulate multiple Hfq chaperones for presentation

Bacterial sRNAs generated with Rho-independent terminators contain a 3' poly-Uridine (poly-U) tail, and these poly-U sequences bind the Hfq proximal side (Schulz & Barabas, 2014). Class II sRNAs also contain an A-rich region towards the 5'-end that can bind the Hfq distal side. Both interaction patterns are present in the PNPase-Hfq-3'ETS^{leuZ} structures described above. Other type II sRNAs, like CyaR, were also predicted to form RNA carrier assemblies with Hfq and PNPase (Cameron & De Lay, 2016; De Lay & Gottesman, 2011). CyaR has several mRNA targets, such as *ompX* and *luxS*, and has been shown to help regulate catabolite repression, quorum sensing, and nitrogen assimilation in *E. coli*. CyaR promotes degradation of *ompX*, which encodes one of the major Outer Membrane Proteins/Porins (OMPs), by directly pairing with it (De Lay & Gottesman, 2009, Johansen *et al.* 2008). In this section, CryoEM was used to probe whether PNPase-Hfq-CyaR RNA carrier assemblies have a similar architecture as the PNPase-Hfq-3'ETS^{leuZ} complex.

A representative cryo-EM image with PNPase-Hfq-CyaR particles is given in Figure 19A. As expected, the KH-S1 portal is conformationally heterogeneous in the PNPase-Hfq-CyaR particles, which is apparent from 2D class averages of the assembly (Figure 19B). 3D reconstructions in cryoSPARC and Relion 3.0 reveal an assembly that resembles the overall PNPase-Hfq-3'ETS^{leuZ} architecture. Only one dataset was collected for PNPase-Hfq-CyaR resulting in reconstructions with a much lower global resolution than the 3'ETS^{leuZ}-based RNA carrier complex. In the final map, at 13Å resolution (GS-FSC), the overall hexameric shape of Hfq was recognisable. Remarkably, the reconstruction shows how the particles contain two stacked Hfq hexamers at the KH-S1 portal (Figure 19C and D). The large distances between the docked Hfq chaperones and the KH-S1 portal domains suggest that CyaR must form the binding interface between PNPase and Hfq, like that seen for the PNPase-Hfq-3'ETS^{leuZ} assembly. Even though the overall resolution was too poor to trace CyaR, these reconstructions indicate that multiple Hfq chaperones can cooperate to present longer class II sRNAs to the PNPase KH-S1 portal to form RNA carrier assemblies. PNPase can accommodate such stoichiometric variability via the intrinsic flexibility of its KH and S1 domains. In the PNPase-Hfq-CyaR carrier complex, for example, one of three engaged S1 domains undergoes significant reorganisation to coordinate the RNA on the rim of the second Hfq (Hfq2) chaperone (Figure 19D). In summary, these results further illustrate the potential structural diversity of ribonucleoprotein complexes with which PNPase can cooperate to control sRNA stability and activity in the cell, and how the flexibly tethered KH-S1 portal enables this.

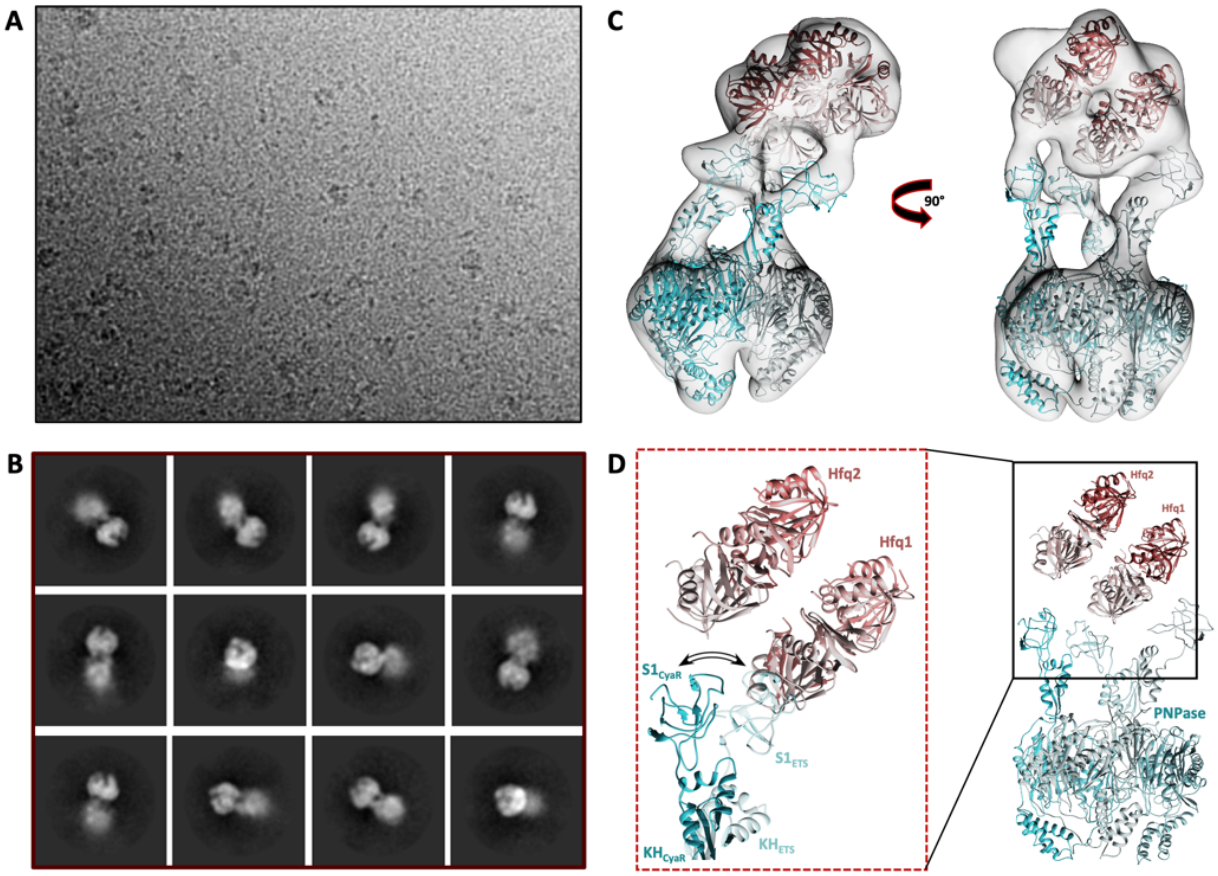


Figure 19: ***CyaR is presented by two Hfq chaperones. (A)*** *Representative cryo-EM image of at 3 μm defocus.* ***(B)*** *2D class averages from cryoSPARC reveal diffuse yet bulky densities at the KH-S1 portal.* ***(C)*** *Careful 3D classification results in a subset of particles with two ring-like densities near the KH-S1 domains, corresponding to two Hfq chaperones.* ***(D)*** *The Hfq hexamers stack on top of each other. The orientation of Hfq2 presented here is speculative due to the limited resolution of the reconstructions. One out of three KH-S1 pairs is subject to a significant outward reorganization to coordinate the Hfq2-RNA pair (inset).*

3.7 Downstream effector roles for PNPase-Hfq-mediated RNA carrier assemblies

The experiments described in section 3.5.1 demonstrate how Hfq and the PNPase KH-S1 portal can cooperate to protect a substrate sRNA from degradation *in vivo* and are in accord with the finding that sRNA species undergo faster turnover in PNPase knock-out mutants (De Lay & Gottesman, 2011; Bandyra *et al.,* 2016). Such protection of the sRNA is necessary in order to guarantee sRNA stability in the cell until it reaches its target. Although the structures above show that PNPase cannot degrade an sRNA in the RNA carrier mode, it is unclear whether these RNA carrier complexes also protect sRNAs from other ribonucleases in the cell. In addition, it is unclear what happens when e.g. an RNA target interacts with an sRNA wrapped in an RNA carrier assembly. In cooperation with Dr. Kasia Bandyra, a series of gel shift- and activity assays were set up to answer these questions. Both experiments were set up in buffer conditions that are non-catalytic for PNPase.

To probe the interaction between a sRNA in an RNA carrier complex and the corresponding RNA target, a series of gel shift assays (EMSA) were set up. The design of the EMSA experiments was such that RNA carrier complexes were assembled prior to adding the respective targets of the sRNAs. The sRNA-target pairs tested were 3'ETS^{leuZ}-RyhB (Figure 20A), CyaR-*ompX* (Figure 20B) and RyhB-*chiP* (Figure 20C). 3'ETS^{leuZ} acts as a sponge for RyhB, also a sRNA, and both sRNAs have been described in section 3.4. Interestingly, both 3'ETS^{leuZ} and RyhB can independently form RNA carrier assemblies with PNPase and Hfq (Bandyra *et al.*, 2016). CyaR and its target *ompX* were described in section 3.6. *ChiP*, which is targeted by RyhB, encodes a conserved enterobacterial chitoporin required for uptake of chitin-derived oligosaccharides (Figueroa-Bossi *et al.*, 2009).

In Figure 20A, a PNPase-Hfq-3'ETS^{leuZ} assembly was preformed (highlighted by *) after which increasing concentrations of RyhB were added. Strikingly, the PNPase-Hfq-3'ETS^{leuZ} complex disassembled even at the lowest RyhB concentrations, resulting in lower order species on the gel. One of these lower order species were 3'ETS^{leuZ}-RyhB pairs (Figure 20, lanes i and ii, red arrow). On the other hand, when mixing 3'ETS^{leuZ}, RyhB and Hfq in the absence of PNPase, a Hfq-RyhB-3'ETS^{LeuZ} complex appeared on the gel, whereas free 3'ETS^{leuZ}-RyhB duplexes were much less abundant (Figure 20A, lane ii and iii). These observations suggest that when 3'ETS^{leuZ} is presented to RyhB by an RNA carrier complex, the main product is a 3'ETS^{leuZ}-RyhB pair, free of Hfq (Figure 20A, lane i, red arrow). It is possible that PNPase-Hfq carrier complexes can present 3'ETS^{leuZ} to RyhB and facilitate release of the RNA duplex from Hfq after sRNA-RNA target pairing. As such, the RNA carrier assembly could organise and expose the RyhB seed region in the 3'ETS^{leuZ} sequence on the Hfq rim (Figure 18).

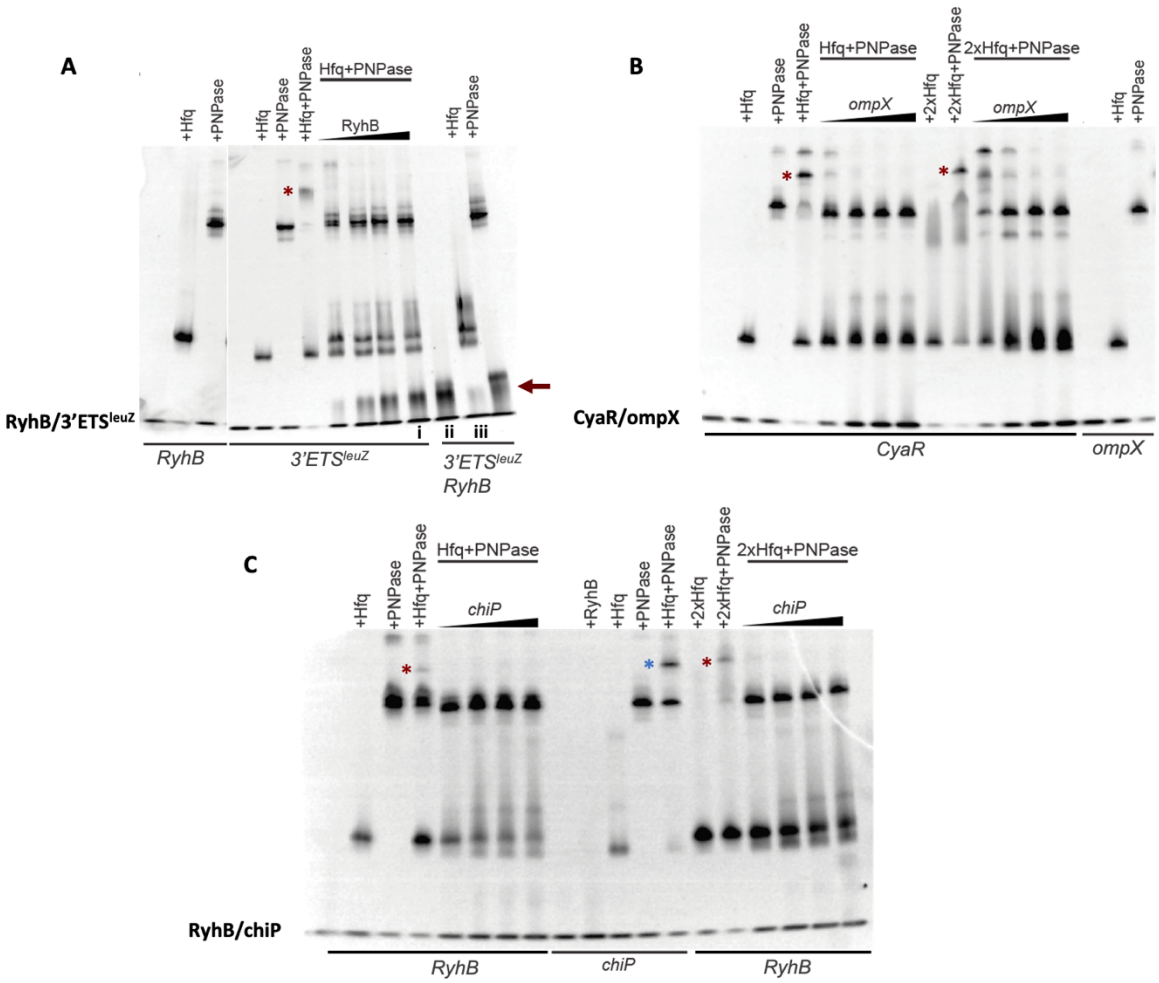


Figure 20: ***RNA carrier assemblies dissociate upon exposure to an RNA target. (A)*** *A stable 3'ETS^{leuZ}-based RNA carrier assembly is dismantled by a titration of RyhB. The RNA carrier assembly facilitates 3'ETS^{leuZ}-RyhB offloading from Hfq. 3'ETS^{leuZ}-RyhB duplexes are highlighted by a red arrow.* ***(B)*** *OmpX can efficiently disband PNPase-Hfq-CyaR ternary complexes, regardless of whether one or two Hfq chaperones per PNPase enzyme are added to the sample. No CyaR-ompX duplexes were observed on the gel.* ***(C)*** *RyhB-mediated RNA carrier assemblies are dissociated when chiP is titrated in, regardless of whether one or two Hfq chaperones are available per RyhB species. No RyhB-chiP duplexes were observed on the gel. RNA carrier assemblies are highlighted by *; non-specific PNPase-RNA-Hfq complexes are highlighted by **.*

Similar results were obtained for the class II sRNA CyaR and a fragment of its target *ompX.* A stable PNPase-Hfq-CyaR complex assembled upon mixing the individual components but dissociated when *ompX* was titrated in, both when one or two Hfq chaperones per PNPase enzyme were added to the sample *(*Figure 20B, ternary complexes are highlighted by *). The rationale for adding a twofold excess of Hfq to the mixture stems from the results discussed in section 3.6, where it is clear that multiple Hfq chaperones can participate in an RNA carrier assembly with longer sRNAs like CyaR. The main products after complex dissociation here were Hfq-RNA species (Figure 20B). Interestingly, free CyaR-*ompX* pairs were not observed when the RNA carrier was disassembled, unlike the 3'ETS^{leuZ}-RyhB pair, and bands for putative Hfq-CyaR-*ompX* complexes were weak. One possible explanation is that PNPase captures nascent CyaR-*ompX* pairs for degradation as these are released from Hfq upon dissociation of the RNA carrier assembly.

Lastly, for the type I sRNA RyhB a ternary complex was visible on the gel albeit at much lower abundance than for the two type II sRNAs tested above (Figure 20C, highlighted by *). Here too the assembly dissociated when the target of RyhB, *chiP*, was titrated in. Overall the gel shift results for RyhB are nearly identical to those for CyaR. Since RyhB is of a similar length as CyaR, the gel shift assays were also performed in the presence of a twofold excess of Hfq, following the observations made in section 3.6. When a twofold molar excess of Hfq was added, a slightly larger ternary complex appeared to be formed, in accord with the results in section 3.6. No RyhB-*chiP* duplex formation was observed on the gel upon complex dissociation by the target. Analogous to CyaR-mediated RNA carrier assemblies, it is possible that PNPase captures nascent RyhB-*chiP* pairs for degradation as these are released from Hfq (Figure 20C). Interestingly, *chiP* also formed a ternary complex with PNPase and Hfq (Figure 20C, highlighted by *). It is likely that that these are non-specific complexes where PNPase captured the 3' end of the mRNA target, which was in turn bound to Hfq. Overall the results described above indicate that RNA carrier complexes could facilitate presentation and handover of sRNAs to their target RNAs, beyond protection of the sRNA from PNPase cleavage. Given that RNA targets can efficiently dissociate RNA carrier assemblies suggests that they directly compete with the PNPase KH and S1 domains for sRNA binding.

Next, experiments were set up to explore whether RNA carrier complexes can protect their sRNA cargo from other ribonucleases in the cell, and whether handover of the sRNA to its target enables or even facilitates cleavage. To this end, activity assays with two key bacterial endoribonucleases, RNase E and RNase III, were carried out. In these assays, RNase E cleavage of RyhB and CyaR was tested when presented by only Hfq or as part of an RNA carrier assembly (Figure 21). RyhB and CyaR stability has previously been shown to be controlled by RNase E (Masse *et al.*, 2003; Kim & Lee, 2020). A truncated version of RNase E (rne1-850) that lacks the PNPase binding site was added as an assembly with two canonical components of the *E. coli* RNA degradosome: the RNA helicase RhlB and the glycolytic enzyme enolase. This assembly is the focus of study in Chapter III and will be referred to as the truncated degradosome. The experiments described here were carried out in buffer conditions that are non-catalytic for PNPase, and its residual degradation activity is negligible. RyhB was not cleaved efficiently by RNase E in the presence of Hfq alone (Figure 21A, 26.0±13.8% of RyhB degraded after 15 minutes), and the formation of a PNPase-mediated RNA carrier complex increased its stability only modestly (14.1±9.6% of RyhB degraded after 15 minutes). On the other hand, RNase E cleaved Hfq-associated CyaR rapidly, suggesting that despite the notion that Hfq alone confers general sRNA stability, some sRNAs are still prone to degradation when associated with the chaperone (Figure 21B, 94.9% of CyaR degraded after 15 minutes). Strikingly, CyaR was protected from RNase E cleavage when wrapped in a PNPase-Hfq-CyaR complex (Figure 21B, 28.5±9.8% of CyaR degraded after 15 minutes).

RyhB has been shown to be degraded by the duplex-cleaving ribonuclease RNase III as well (Afonyushkin *et al.*, 2005), and activity assays were set up to test whether RyhB is protected from RNase III cleavage when part of an RNA carrier assembly. The presence of PNPase in addition to Hfq resulted in an increase of RyhB stability compared to the reaction with only Hfq (Figure 21C left panel, 38.7±5.6% and 52.2±2.4% RNA degraded after 15 minutes respectively). 3'ETS^{leuZ} on the other hand was not measurably degraded by RNase III, whether presented by Hfq or a RNA carrier

assembly (Figure 21C, middle panel). The gel shift assay presented in Figure 20A indicates that a Hfq-free RyhB-3'ETS^{leuZ} duplex is the main product upon dissociation of the 3'ETS^{leuZ} RNA carrier assembly by addition of RyhB. The base-pairing region between these two RNAs spans 20 nucleotides (Lalaouna *et al.*, 2015) (Figure 18), which makes the RNA duplex a good candidate substrate for RNase III. Therefore, RyhB degradation by RNase III when adding the 3'ETS^{leuZ} sponge to the reaction mixture was tested (Figure 21C, right panel). Notably, RNase III cleaved the RyhB-3'ETS^{leuZ} pairs efficiently. When originating from the RyhB RNA carrier assembly, RyhB-3'ETS^{leuZ} degradation was faster than when only Hfq was present in the sample (88.6±2.1% and 78.2±3.7% of RyhB degraded after 15 minutes, respectively). The results in this section indicate that PNPase confers sRNA protection from degradation by other ribonucleases, but could also increase the efficiency of subsequent RNA degradation in the presence of a base pairing RNA partner. This indicates again that PNPase may boost the facilitating role of Hfq to pair sRNAs with their corresponding target RNAs.

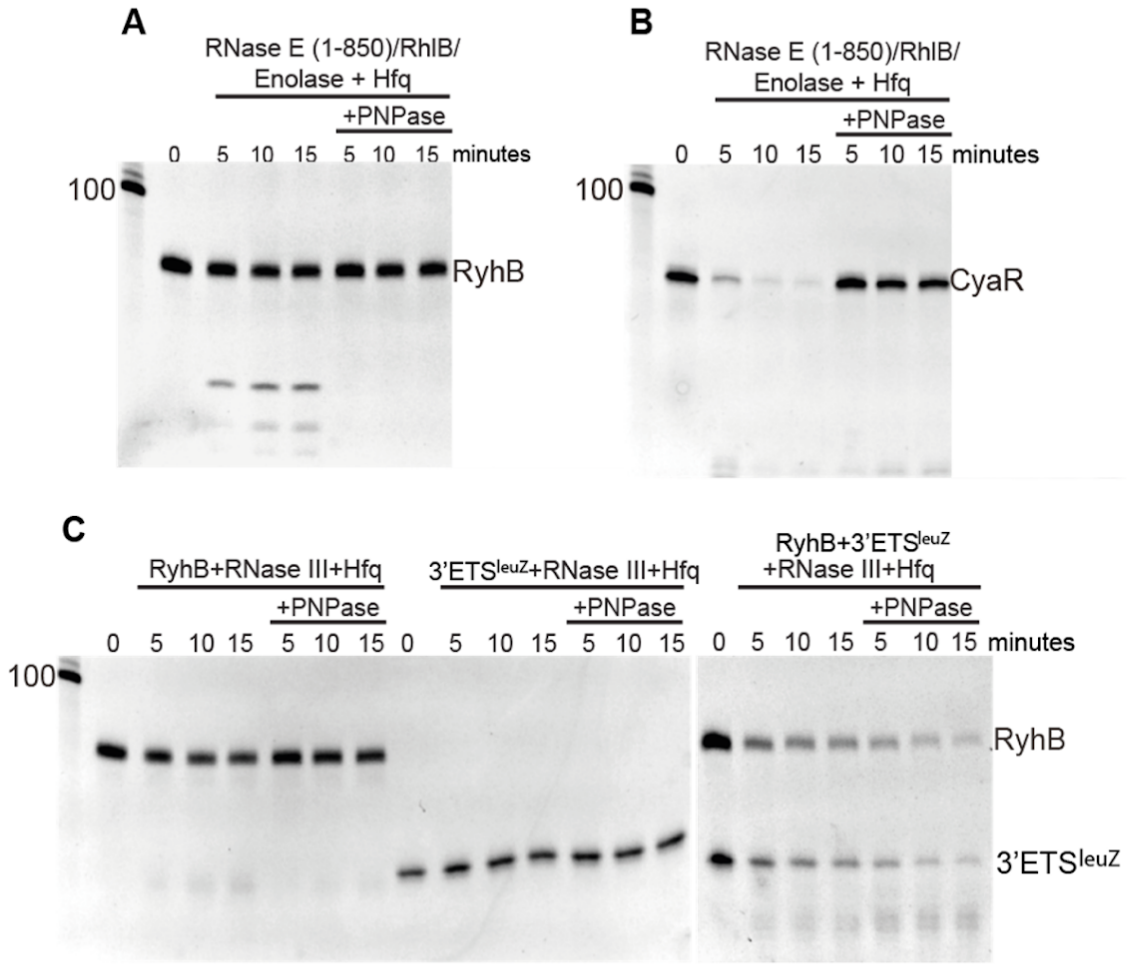


Figure 21: ***RNA carrier assemblies protect sRNAs from RNase E and RNase III.*** *Degradation of RyhB* ***(A)****, CyaR* ***(B)*** *and RyhB/3'ETS^{leuZ}* ***(C)*** *by RNase E* ***(A, B)*** *and RNase III* ***(C)****, in the absence and presence of PNPase. For the RNase E experiments, 200 nM RNA was incubated with 200 nM Hfq and 7.5 nM RNase E (1-850)/RhlB/Enolase in the presence or absence of 200 nM PNPase. For the RNase III experiments, 200 nM RNA was incubated with 200 nM Hfq and 0.01 U (individual RNAs) or 0.005 U (RyhB and 3'ETS^{leuZ} together) RNase III, in the presence or absence of 200 nM PNPase. Reaction times, in minutes, are given in the lanes above the gels. The first lane in each gel is a marker corresponding to 100 nucleotides.*

4 Discussion

The first electron microscopy studies of PNPase were performed more than half a century ago (Valentine *et al.*, 1969). Torus shaped particles were visible on EM micrographs after heavy metal coating of the PNPase particles. In this chapter, 50 years later, the same technique was used to elucidate the different modes of action of PNPase after extensive optimisations of sample preparation and grid freezing. The structure of *E. coli* PNPase in apo-form reveals for the first time the organisation of the KH-S1 portal. Through symmetry expansion and subsequent masked 3D classification, different states of the KH and S1 domains were resolved, showing the conformational versatility of the PNPase RNA 'recognition module' (Figure 7 & Figure 22). Interestingly, Gholzarroshan *et al.* (2018) have postulated that the S1 domains could come together based on SAXS (small angle X-ray scattering) studies of hPNPase (human PNPase). Their ensemble models are strikingly similar to the synchronised sub-states depicted in Figure 7C. The cryo-EM maps presented here, however, indicate that these KH-S1 sub-states are unlikely to be fully synchronised, otherwise they would average better. These results suggest that the conformationally heterogeneous KH-S1 portal could constitute a miscellaneous interaction platform that enables PNPase to fulfil a wide range of cellular roles. The observations described in this chapter certainly support this hypothesis.

The best studied role of PNPase is that of an exoribonuclease, i.e. the enzyme operating in phosphorolytic mode. The structure of the 3'ETS^{leuZ}-bound form of PNPase reveals how the S1 and KH domains of the enzyme come together to intercept single stranded RNA substrates (Figure 8 & Figure 22). The 3'-end of the substrate is then directed into the central channel of the RNase PH core, where the RNA is digested by phosphorolysis. Interestingly, the way in which the KH and S1 domains compact around RNA substrates can also be seen in the transitions between some of the sub-states of the apo-PNPase structure (Figure 7B and Figure 8). These observations support the hypothesis that the loose 'beads on a string'-like organisation of the KH and S1 domains enable PNPase to efficiently capture RNA targets for degradation. It is possible that a different active site is used for every phosphorolytic attack of a phosphodiester bond, which would hint towards a cooperative mode of action between the active sites in the RNase PH domains. The structures for *E. coli* apo-PNPase and PNPase engaged on RNA substrates are the most complete to date, and the first cryo-EM models of any PNPase enzyme.

Since the discovery of PNPase by Severo Ochoa and Marianne Grunberg-Manago more than sixty years ago, its main contribution to bacterial gene regulation was believed to be RNA processing and degradation, acting as described in the previous paragraph for the phosphorolytic mode (Grunberg-Manago & Ochoa, 1955). Over the last decade, research has uncovered an important additional role for PNPase in regulation of sRNA function and stability (De Lay & Gottesman, 2011). It was found that in *E. coli* and *Salmonella Typhimurium,* association with Hfq determines whether some, but not all, sRNAs are degraded or stabilised by PNPase. sRNAs that don't associate with Hfq are usually degraded by PNPase after initial cleavage by RNase E (Viegas *et al.*, 2007; Soderbom & Wagner. 1998; Andrade *et al.*, 2012). For Hfq-dependant sRNAs, such as type I sRNAs RyhB and GcvB and type II sRNAs CyaR and MgrR, PNPase will degrade the sRNA in the absence of Hfq but stabilises them when presented by Hfq (Bandyra *et al.*, 2016; Cameron & De Lay, 2016). The paradoxal observation that sRNAs have reduced stability

in PNPase deletion strains support the hypothesis of additional regulatory roles for PNPase beyond RNA turnover (Bandyra *et al.* 2016; De Lay & Gottesman 2011; Andrade *et al.* 2012*;* Cameron & De Lay, 2016). How PNPase is repurposed by Hfq was captured in this chapter via cryo-EM structures of an RNA carrier assembly involving PNPase, Hfq and a type II sRNA sponge, 3'ETS^{leuZ}. In particular, the structural data reveal how Hfq and the PNPase KH-S1 portal cooperate to coordinate an sRNA and direct it away from the catalytic PNPase core (Figure 14 & Figure 22). All three KH domains and all three S1 domains bind 3'ETS^{leuZ}. *In vitro* binding assays confirm that a series of basic residues on the KH and S1 domains are crucial for complex formation (Figure 16). Interestingly, Hfq presents 3'ETS^{leuZ} so that only the PNPase RNA recognition module (KH-S1) is engaged. The ability of PNPase to lend its RNA recognition module to other effector complexes is likely to be the basis for its pleiotropic roles in posttranscriptional regulation (Cameron *et al.*, 2018).

Remarkably, the RNA carrier structure shows for the first time how a class II sRNA interacts with Hfq. That is, a 5' A-rich region engages the Hfq distal side, the RNA then wraps around the circumferential rim for the 3' poly-uridine tail to bind the Hfq proximal side (Figure 17). In the PNPase-Hfq-3'ETS^{leuZ} RNA carrier assembly it is mainly the RNA on the Hfq distal side that interacts with the KH-S1 portal, where it forms an incomplete Brennan crown (Figure 17). This is in agreement with the 3'ETS^{leuZ} sequence, which does not contain six ARN-triplet repeats. Instead, only three ARN-triplet repeats could be modelled in the cryo-EM map. Interestingly, however, all six R-sites on the Hfq distal sides are occupied, whereas three A-sites are not occupied. As such, the PNPase-Hfq-3'ETS^{leuZ} structure shows how an imperfect ARN-triplet repeat motif in the 3'ETS^{leuZ} sequence, 5'-RN ARN ARN ARN RNN NR-3', is still presented to a partner effector molecule by the Hfq distal side (Figure 17) (A is Adenine, R is a purine, N is any nucleotide). In light of the molecular dynamics simulations on the Hfq-Crc complexes presented in Chapter I, it is likely that the binding mode between the Hfq distal side and a target sRNA is transient but can be stabilised via interactions with the PNPase KH and S1 domains. For example, due to partly unoccupied A-sites on the Hfq distal site, the RNA backbone lifts off the Hfq distal side and is caught by a PNPase S1 domain (Figure 17C, left panel), which could help stabilise the interaction.

For longer type II sRNAs, like CyaR, cryo-EM reconstructions show how two Hfq chaperones can participate in the formation of an RNA carrier assembly. Remarkably, PNPase can accommodate a second Hfq hexamer via a reorganisation of one of its KH-S1 pairs to adopt a more open KH-S1 portal (Figure 19D). An open question is why two Hfq hexamers would be needed to present CyaR to PNPase, and whether more than two Hfq hexamers would be needed for even longer sRNA substrates. Interestingly, class I sRNAs, such as RyhB, are also stabilised by formation of ternary complexes with PNPase and Hfq (Bandyra *et al.*, 2016; Cameron and De Lay, 2016; De Lay and Gottesman, 2011). Since class I sRNAs are not predicted to bind the Hfq-distal side, it is unclear how Hfq would present such RNAs to PNPase.

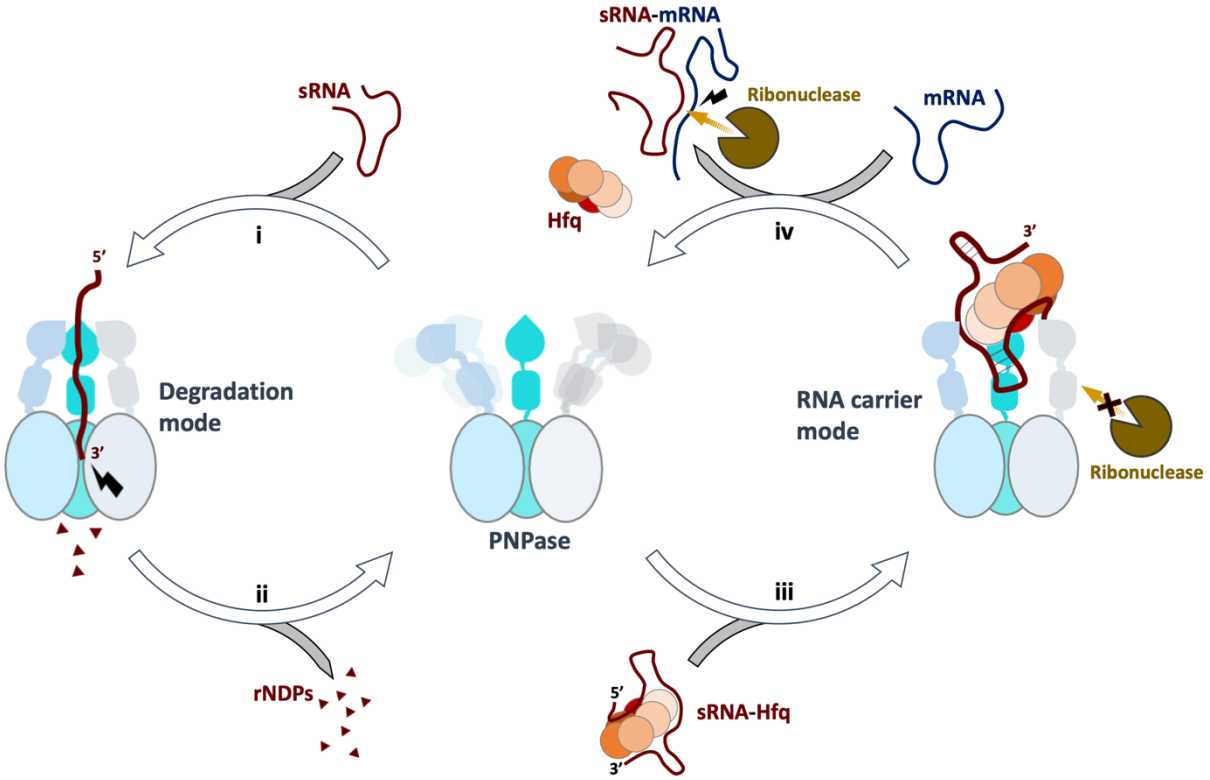


Figure 22: ***Schematic overview of the RNA degradation and RNA carrier modes of PNPase.*** *Apo-PNPase can capture a sRNA via its flexibly tethered KH-S1 portal for degradation in the core via phosphorolysis (i), releasing nucleoside diphosphates (ii). When the sRNA is presented by Hfq, PNPase switches to its RNA carrier mode, lending its KH-S1 portal as a binding platform (iii). Wrapped in an RNA carrier assembly, the sRNA is protected against other ribonucleases in the cell, such as RNase E and RNase III. If an RNA target engages the RNA carrier complex the sRNA is released to pair with the target, which results in dismantling the PNPase-Hfq ternary complex (iv). Ribonucleases, including PNPase can now cleave the RNA target.*

The structural data and complementary cell-based mutagenesis assays describe how RNA carrier assemblies form and confirm their relevance *in vivo*. However, these experiments do not elucidate the functional roles of repurposed PNPase assemblies beyond protection of the sRNA from PNPase activity itself. A series of EMSAs and activity assays resulted in an interesting set of putative roles for PNPase-Hfq regulatory complexes. EMSAs show PNPase-Hfq based RNA carrier assemblies form on 3'ETS^{leuZ}, CyaR (both type II sRNAs) but also on RyhB, a type I sRNA. Strikingly, when the target RNAs of these respective sRNAs are added to the mixture, the RNA carrier complexes disassemble. In the case of 3'ETS^{leuZ} and its target RyhB, one of the main products after the ternary complex is dismantled are Hfq-free 3'ETS^{leuZ}-RyhB pairs. These findings suggest that PNPase-Hfq carrier assemblies could facilitate pairing of 3'ETS^{leuZ} to its RNA target RyhB, and subsequent offloading from Hfq (Figure 20A). CyaR- and RyhB-based RNA carrier complexes are also dismantled upon addition of their RNA targets, yet no Hfq-free sRNA-RNA target duplexes were observed on the gels. It is possible, however, that these duplexes are still bound to PNPase. Presumably, many cellular RNAs will compete with a sRNA for PNPase or Hfq binding and can disrupt the RNA carrier assembly. Therefore, off-target dismantling of the RNA carrier assembly is likely to be prevalent in the cell, without downstream RNA pairing. Lastly, activity assays reveal how the interactions in the ternary complex can shield CyaR and RyhB from digestion by RNase E. In addition, RNA carrier assemblies can shield RyhB from RNase III digestion. Remarkably, when adding 3'ETS^{leuZ} to the reaction mixture, RNase III digestion of the

putative RyhB-3'ETS^{leuZ} pairs is very efficient. In summary, it is likely that RNA carrier assemblies can protect sRNAs from degradation in the cell and help present them to their target mRNAs to facilitate subsequent regulation. With the ternary complex structures in hand, I sought to elucidate the structural basis for the apparent facilitating role of PNPase on sRNA action by imaging quaternary complexes with the target RNAs (e.g. PNPase-Hfq-3'ETS^{leuZ}-RyhB). However, extensive efforts to make the quaternary assembly with target mRNAs resulted only in PNPase-RNA particles, and the results in Figure 20 may offer an explanation, in that the ternary complex is inherently destabilised by target RNAs.

In *E. coli* and other Gram-negative bacteria, Hfq-bound sRNAs regulate many cellular processes, such as biofilm formation, antibiotic resistance, DNA repair and motility (Holmqvist *et al.* 2010; Jorgensen *et al.* 2013; Nishino *et al. 2011;* Thomason *et al.*, 2012; Chen & Gottesman, 2017). Given this novel role of PNPase in sRNA mediated regulation, many of the phenotypes associated with loss of PNPase are potentially explained by its dual role in sRNA stability and presentation. Modulation of PNPase activity is not limited to *E. coli* and RNA carrier complexes built on PNPase have potentially evolved beyond bacteria. In humans, PNPase (hPNPase) localises in between the mitochondrial inner and outer membrane (Inter Membrane Space, IMS) (Chen *et al.,* 2006). Many roles have been reported for hPNPase, including mitochondrial RNA import. In particular, it has been shown that hPNPase binds 5S rRNA and RNase MRP/RNase P RNAs to help mediate import into the mitochondrial matrix (Wang *et al.*, 2010). Strikingly, catalytic activity is not necessary for PNPase to facilitate transport, and a short, 20-nucleotide stem-loop in the RNase P and RNase MRP RNAs serves as an anchor for PNPase. Golzarroshan *et al.* further showed that the S1 domains of hPNPase are crucial for binding these stem-loops, whereas they are not needed for binding and degrading single stranded sRNA species (Golzarroshan *et al.*, 2018). This is in agreement with the structures above, where a short stem-loop in 3'ETS^{leuZ} is extensively coordinated by the *E. coli* KH-S1 portal. Through its role in facilitating RNase P and RNase MRP RNA transport across the mitochondrial IMS, PNPase directly controls mitochondrial RNA maintenance and cellular respiration (Shimada *et al. 2018;* Shokolenko & Alexeyev, 2017). Besides 5S rRNA and RNase MRP/RNase P RNAs, PNPase was also shown to facilitate the import of miRNAs, such as miR-378, into mitochondria (Shepherd *et al.*, 2017). Finally, and perhaps more controversial, hPNPase is believed to mediate the import of long non-coding RNAs such as the RNA component of telomerase (TER) for processing, after which it is again exported to the cytoplasm by hPNPase (Cheng *et al.*, 2018). Clinical studies show that some phenotypes related to mutations in *PNPT1*, the hPNPase gene, such as axonal neuropathy, Leigh Syndrome, hereditary hearing loss and delayed myelination could be related to impaired mitochondrial RNA import (Sato *et al.*, 2018; Vedrenne *et al.*, 2012; von Ameln *et al.*, 2012). How hPNPase mediates transport across the IMS remains to be answered, but the KH-S1 portal is likely to play a crucial role, and there might be similarities with the way RNA carrier complexes in *E. coli* lend their KH and S1 domains to other effector molecules.

The structures above display new interactions between putative RNA binding motifs in conserved domains and target RNAs. For example, the structure of the ternary complex reveals a mode of cooperation between an S1 domain and Hfq to interact with RNA. In *E. coli*, the S1 protein of the ribosomal small subunit has a string of four S1 RNA binding modules that promote accommodation into the decoding channel of the 30S subunit (Duval *et al.*,

2013), and these might form similar interactions with transcripts bound by Hfq. A recent study on 100S hibernating ribosomes revealed how one of the S1 domains interacts with the anti-Shine-Dalgarno sequence of the 16S rRNA (Beckert *et al.*, 2018). This interaction pattern is similar to the interactions observed for the PNPase-S1 domain and 3'ETS^{leuZ} in the RNA carrier assembly (S1-1, Figure 14D). Therefore, this model might explain how the Hfq distal side and ribosomal S1 could act together on certain transcripts to help determine their fate. In this mode, the proximal face of Hfq would be exposed to recruit sRNAs that can remodel the translation machinery to either support translation or to trigger degradation by ribonucleases such as RNase E.

In Summary, the results presented in this chapter elucidate an alternative function for bacterial PNPase in the cell. PNPase does not solely degrade or process RNA substrates but can lend its KH-S1 portal as a complex interaction platform for sRNAs and their RNA targets. Repurposing of PNPase into this RNA carrier mode is triggered by Hfq-mediated presentation of a sRNA. As long as the sRNA is wrapped in the RNA carrier assembly, it is protected from ribonucleases such as RNase E/the RNA degradosome and RNase III. When a target RNA engages the RNA carrier complex it dissociates, releasing Hfq. After handover, the sRNA-RNA target pair can be targeted by ribonucleases. Given the many key functions of sRNAs in bacteria, it is clear that PNPase can act as a nuanced but pivotal node in post-transcriptional regulation networks.

5 Materials and methods

5.1 Hfq expression and purification

pEH-10-(hfq) was transformed into chemo-competent BL21(DE3) cells via standard heat shock transformation. 2×YT media (Formedium) supplemented with 100 µg/ml carbenicillin was inoculated with overnight precultures and grown at 37°C. The cultures were induced with 1 mM IPTG at OD_{600nm} 0.45. 3 hours after induction, cells were harvested by centrifugation (Beckman 4.2 rotor, 5000 g; 20 minutes), resuspended in lysis buffer (50 mM Tris pH 8, 1.5 M NaCl, 250 mM MgCl2, 1 mM EDTA, 1 protease inhibitor cocktail tablet (Roche)) and flash frozen in liquid nitrogen.

Upon thawing the cells were passed thrice through an EmulsiFlex-05 cell disruptor for lysis (10-15 kbar, Avestin). The lysate was clarified with centrifugation (4°C, 30 minutes, 37500 g) and incubated in an 85°C water bath for 20 minutes. The lysate was then centrifuged again (20°C, 30 minutes, 37500 g) to pellet the precipitate. Ammonium sulfate was added to the supernatant up to 1.6 M followed by another centrifugation step (20°C, 30 minutes, 37500 g). The final supernatant was loaded onto a HiTrap Butyl–Sepharose column (GE Healthcare) equilibrated with HfqA buffer (50 mM Tris pH 8.0, 1.5 M NaCl, 1.5 M (NH4)2SO4). Proteins were eluted with an isocratic gradient of HfqB buffer (50 mM Tris pH 8.0) and analysed on SDS–PAGE. Fractions enriched with Hfq were pooled, concentrated to 2 ml with a 15 ml Amicon Ultra 30,000 MWCO concentrator (Millipore) and loaded onto a Sephadex 75 gel filtration column (GE Healthcare) equilibrated with HhfqC buffer, containing 50 mM Tris pH 8.0, 100 mM NaCl, 100 mM KCl and 1 protease inhibitor cocktail tablet (Roche). Fractions were analysed by SDS– PAGE and the concentration of clean Hfq fractions was determined spectroscopically using a NanoDrop ND-1000 spectrophotometer (Thermo Scientific) and a λ280nm extinction coefficient of 4470 $M^{-1}cm^{-1}$ (per Hfq protomer) before the samples were flash frozen and stored at -80°C. Each fraction was also analysed via SDS-PAGE (Figure 3)

5.2 PNPase expression and purification

pET_Duet_pnpwt was transformed into chemo competent BL21(DE3) cells via standard heat shock transformation. 2×YT media (Formedium) supplemented with 100 µg/ml carbenicillin was inoculated with precultures (grown for 2hr at 37°) and grown at 37°C. The cultures were induced with 0.5 mM IPTG at OD_{600nm} 0.45 and the temperature was reduced to 25°C. 3-4 hours after induction, cells were harvested by centrifugation (5000 g, 20 minutes), resuspended in lysis buffer (20 mM Tris pH 8, 150 M NaCl, 150 mM KCl, 5 mM MgCl2, 1 mM EDTA, 1 protease inhibitor cocktail tablet (Roche)) and flash frozen in liquid nitrogen.

Upon thawing the cells were passed thrice through an EmulsiFlex-05 cell disruptor for lysis (10-15 kbar, Avestin). After lysis, the lysate was clarified by centrifugation (4°C, 30 minutes, 37500 g) and PNPase was precipitated with ammonium sulphate (51.3% saturation) for 1 hour at 4°C. The sample was centrifuged (4°C, 30 minutes, 37500 *g*), and the pellet was resuspended in Q Buffer A' (20 mM Tris-HCl pH 8.5, 0.5 mM TCEP, 10% v/v glycerol, EDTA-free protease inhibitor cocktail (Roche)), loaded on 5 ml HiTrap Q column (GE Healthcare) equilibrated in Q buffer A (20

mM Tris-HCl pH 8.5, 30 mM NaCl, 0.5 mM TCEP, 10% v/v glycerol) and eluted with a 0-60% gradient of Q Buffer B (Q Buffer A with 1 M NaCl). Fractions from the Q column were evaluated by SDS-PAGE and those containing PNPase were pooled and supplemented with 1mM MgCl2 and 45 mM Na phosphate pH 7.9 and left on the bench for 30 min to let PNPase degrade the remaining RNAs. Next, 0.9 M $(NH_4)_2SO_4$ and 1 mM TCEP were added and the protein solution was loaded on a 5 ml HiTrap Butyl-Sepharose column (GE Healthcare) equilibrated with BS Buffer A (50 mM Tris-HCl pH 7.5, 1 M $(NH_4)_2SO_4$, 0.5 mM TCEP) and eluted with a 0-66% gradient of BS Buffer B (50 mM Tris-HCl pH 7.5). Fractions containing PNPase were pooled, concentrated with Amicon Ultra 30,000 MWCO filter units (Millipore) and loaded on a Superdex 200 10/300 GL gel filtration column (GE Healthcare) equilibrated with a buffer composed of 20 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM $MgCl_2$, 0.5 mM TCEP, 10% (v/v) Glycerol. Purified protein concentration was determined spectroscopically using a NanoDrop ND-1000 spectrophotometer (Thermo Scientific) and a λ_{280nm} extinction coefficient of 30370 $M^{-1}cm^{-1}$ (per PNPase protomer). Protein fractions were flash frozen in liquid nitrogen and stored at -80°C. All peak fractions were analysed via SDS-PAGE (Figure 3). All PNPase mutants used in this study were cloned into pET_Duet and expressed/purified in a NRD1617 Δpnp strain (BL21(DE3)) via the protocol described above. The cloning entails is a tailored three step PCR-fusion protocol and will be described in detail in Chapter III.

5.3 *In vitro* transcription and RNA purification

All RNAs used in this study (3'ETS^{leuZ}, RyhB, CyaR, *ompX*, *chiP*), except for GcvB, were prepared by in vitro transcription (IVT) according to standard protocol. The templates for IVT for RyhB and CyaR were PCR products amplified from a vector carrying the target sequence (available in the lab) with Phire Hotstart II polymerase (Thermo Fisher) (Table 3). The templates for IVT for 3'ETS^{leuZ}, *chiP* and *ompX* were generated by heating overlapping oligomers (Sigma) at 95°C for 15 minutes and cooling them down to room temperature by turning off the heat block for hybridisation (Table 3, highlighted with *). The sequences for PCR primers and DNA oligonucleotides are summarised in Table 3. Both the primers and oligonucleotides contain a promoter sequence recognized by T7 RNA polymerase during IVT. IVT was run for 5 hours at 37°C and quenched with Turbo DNase I (Ambion) and 50 mM EDTA (final concentration). IVT products were purified on a 6-8% polyacrylamide denaturing gel with 7M Urea (National Diagnostics), bands containing RNA were visualised by UV-shadowing, excised, electroeluted in TBE (Whatman Elutrap) (overnight) and cleaned up using PureLinkTM RNA Microscale Kit (Invitrogen). CyaR samples were prepared by Dr. Kasia Bandyra, Luisi group, Department of Biochemistry, University of Cambridge.

Table 3: Primers and oligonucleotides (*) used for preparation of IVT templates

Primer	Sequence (5'>3')
RyhB_F	TATAGAATTCTAATACGACTCACTATAGCGATCAGGAAGACCCTC
RyhB_R	AAAAAAAAAGCCAGCACCCGGCTGGCTAAG
ETS_F*	GTTTTTTTAATACGACTCACTATAGGGAAAGATAAGAATAAAATCAAAGCAATAAGCAGTGTCGTGAAA CCACCTTCGGGTGGTTTTTTT

ETS_R*	AAAAAAACCACCCGAAGGTGGTTTCACGACACTGCTTATTGCTTTGATTTTATTCTTATCTTTCCCTATAG TGAGTCGTATTAAAAAAAC
GcvB_F	GTTTTTTTTTTAATACGACTCACTATAGACTTCCTGAGCCGGAACGAAAAG
GcvB_R	AAAAAAAAGCACCGCAATTAGGCGGTG
CyaR_F*	GTTTTTTTTTTAATACGACTCACTATAGCTGAAAAACATAACCCATAAAATGC
CyaR_R*	AAAAAATAAGCCCGTGTAAGGGAGATTACACAGG
chiP_F*	GTTTTTTAATACGACTCACTATAGGTTTTTATTTTTTCGCTGTTCACCTTTGGTGCAGCAATTTATACGTCA AAGAGGATTAAC
chiP_R*	GTTAATCCTCTTTGACGTATAAATTGCTGCACCAAAGGTGAACAGCGAAAAAATAAAAACCTATAGTGA GTCGTATTAAAAAAC
ompX_F*	GTTTTTTAATACGACTCACTATAGGACTTATTTGAATCACATTTGAGGTGGTTatgAAAAAAATTGCATGT CTTTCAGCACTG
ompX_R*	CAGTGCTGAAAGACATGCAATTTTTTTcatAACCACCTCAAATGTGATTCAAATAAGTCCTATAGTGAGTC GTATTAAAAAAC

5.4 Denaturing gel-electrophoresis

Denaturing gel-electrophoresis of RNA samples was performed at room temperature using polyacrylamide gels with 7 M Urea, in 1× TBE buffer (SequaGel UreaGel System, National Diagnostics). Samples were mixed with RNA loading dye (Thermo Fisher) and heated to 95°C for 3 minutes before loading them on the gels. Gels were run for 30 min at 200V or for 3 – 5 hours at 240-260 V depending on the size of the gel. Gels were stained for 30-60 min in SYBR Gold solution (Invitrogen) and RNA was visualised with GeneSnap.

5.5 PNPase cloning

PNPase mutagenesis was carried out by Dr. Kasia Bandyra, Luisi group, Department of Biochemistry, University of Cambridge. Mutants of the PNPase gene for activity- and gel-shift assays were prepared by two successive PCR reactions. In the first step, two fragments were PCR-amplified using the wild type PNPase gene as a template: one using the forward primer PNPaseNcoFor and a reverse primer introducing the mutation (primers PNPaseS1x2Rev or PNPaseKHx2Rev or PNPaseS1x4Rev) (Table 4); the other using a reverse primer PNPaseNotRev and a forward primer introducing the mutation (primers PNPaseS1x2For or PNPaseKHx2For or PNPaseS1x4For) (Table 4). PCR products were resolved on 1% low melting point agarose gel (Sigma), the bands of interest were excised and after melting the matrix at 70°C, they were mixed in one PCR reaction which amplified the entire PNPase gene, with mutations, using primers PNPaseNcoFor and PNPaseNotRev. The product of the last PCR was digested with NcoI and NotI (NEB), resolved on a low melting point agarose gel, and the gel band was directly ligated with T4 ligase (NEB) into a pET duet plasmid, which had been digested with the same restriction enzymes and dephosphorylated with CIP (NEB) according to the manufacturer's instructions. A detailed description of this cloning strategy will be described in Chapter III.

Table 4: Primers for PNPase cloning

Primer	Sequence (5'>3')
PNPaseNcoFor	TATACCATGGTTGCTTAATCCGATCGTTCGTAAATTCC
PNPaseNotRev	TATATAGCGGCCGCTTACTCGCCCTGTTCAGCAGCCG
PNPaseS1x2For	CTCTCAAATCGCTGACGCGGCGGTTGAGAAAGTGACCG
PNPaseS1x2Rev	CGGTCACTTTCTCAACCGCCGCGTCAGCGATTTGAGAG
PNPaseKHx2For	CAACCCGGACAAGATCGCGGATGTTATCGGTGCGGGCGGTTCTGTAATCC
PNPaseKHx2Rev	GGATTACAGAACCGCCCGCACCGATAACATCCGCGATCTTGTCCGGGTTG
PNPaseS1x4For	GTTCTGGAAGTTGATGCGGCGGGCGCGATCGCGCTGAGCATTAAAGAAG
PNPaseS1x4Rev	CTTCTTTAATGCTCAGCGCGATCGCGCCCGCCGCATCAACTTCCAGAAC

5.6 Cryo-EM sample preparation

5.6.1 Complex stabilisation with GraFix

A continuous 0-30% (v/v) glycerol gradient was set up in 4.5ml tubes with a Gradient Master 108 (Biocomp instruments), according to standard protocols. In parallel, a dual 0-30% (v/v) glycerol/0-0.15% (w/v) glutaraldehyde gradient was made. The base buffer for both was 20 mM Hepes pH 8.0, 150 mM NaCl, 5 mM $MgCl_2$, 0.5 mM TCEP. After dialysis of the samples in the base buffer, 200 nmoles of PNPase, PNPase-sRNA or a preformed PNPase-Hfq-3'ETS^{leuZ} complex were added on top of the gradients. The tubes were subjected to ultracentrifugation (SW60-Ti) for 18 hours at 164000 g.

Gradients were fractionated into 200 microL fractions with a peristaltic pump (at 4°C) and the tube with crosslinker was quenched with 10mM Tris-HCl pH7.5. Fractions coming from the gradient without glutaraldehyde were precipitated by adding 10% (v/v) of trichloroacetic acid (50% stock). Precipitate was pelleted by 30 minutes centrifugation at 15000 g (4°C) and subsequently resuspended in SDS-PAGE loading dye (TruPage LDS sample buffer, 200mM DTT). The protein content was analysed via SDS-PAGE. For the fractions containing the assembly of interest, the corresponding fractions in the glutaraldehyde-glycerol gradient were pooled, dialysed against the GraFix base buffer (20 mM Hepes pH 8.0, 150 mM NaCl, 5 mM $MgCl_2$, 0.5 mM TCEP) to remove glycerol and concentrated down to 50 µl. CryoEM grids were prepared with an FEI Vitrobot (IV) and screened on a FEI Talos Arctica.

5.6.2 Preparation cryo-EM samples supplemented with CHAPSO

PNPase, PNPase-sRNA and PNPase-Hfq-sRNA samples were prepared at 8-12 µM in a buffer containing 20 mM Tris-HCl pH 8, 2.5 mM $MgCl_2$, 150 mM KCl and 1 mM TCEP. 3 µl of each sample was applied to glow discharged (EasiGlow Pelco) R2/2 Au Ultrafoil grids (Quantifoil), immediately after addition of 8 mM CHAPSO (3-([3-cholamidopropyl]dimethylammonio)-2-hydroxy-1-propanesulfonate). Excess sample was blotted away with an FEI Vitrobot (IV) (100% humidity, 4 °C, blotting force {-2,+2}, 3s blot time). After blotting, the grids were vitrified in liquid ethane and screened on a FEI Talos Arctica (BioCem, Department of Biochemistry, University of Cambridge).

5.7 CryoEM data collection and image processing

All datasets were collected on a FEI Titan Krios with a Gatan K2 camera. For PNPase-Hfq-3'ETS^{leuZ}, four separate datasets were collected, of which two with a Volta Phase Plate (VPP, FEI). For the PNPase-sRNA constructs, all processing was performed in Relion 3.0 (Scheres, 2012; Zivanov *et al.*, 2018). For the apo-PNPase and PNPase-Hfq-3'ETS^{leuZ} datasets, processing was done in Relion 3.0, after which polished particles were exported to cryoSPARC (Punjani *et al*., 2017) for further 3D classification and refinement rounds as well as 3D variability analysis. For all datasets, 2D templates were generated for autopicking by manually picking 1000-2000 particles and running 2D classification on them. Several 2D and 3D classification rounds were used to clean the picked set of particles before refining the maps. Typically, after performing per-particle CTF estimation and per-particle motion correction/per frame radiation-damage weighting another round of 3D refinement was set up. The 3D auto-refinements resulted in consensus maps at intermediate- to high resolutions (Figure 6 and Figure 11). From here a tailored approach was used to further process the apo-PNPase and PNPase-Hfq-3'ETS^{leuZ} particle sets to deal with conformational heterogeneity and increase the information content in the maps. These approaches are discussed in the next two paragraphs.

To better resolve the flexible KH-S1 portal in the apo-PNPase reconstruction, a consensus refinement was set up with C3 symmetry. Next, the symmetry operators were used to perform symmetry expansion on the particle set. A wide mask was generated for a PNPase protomer containing the KH and S1 domains and a masked 3D classification was run without alignments, to prevent overfitting. After two more rounds iterative mask improvement and subsequent 3D classification, the KH-S1 portal was much better resolved. In addition, three different conformational states of the KH and S1 domains were found (Figure 7). In parallel polished particles were transferred to cryoSPARC 2.15 and subjected to symmetry expansion and subsequent 3D variability analysis (3DVA) (6 modes, 6Å resolution limit). (Figure 6) (Punjani & Fleet, 2020).

The PNPase-Hfq-3'ETS^{leuZ} datasets were processed in Relion 3.0 as described above. Due to significant conformational heterogeneity of the KH-S1 portal region compared to the PNPase catalytic core, careful 3D classification (both global and focused) was necessary to resolve the hexameric Hfq chaperone in the complex as an interpretable feature. Multibody refinement and principal component analysis indicated that there was a continuous series of sub-states for the KH-S1-3'ETS^{leuZ}-HFQ region (130 kDa in size) of the complex. However, due to internal domain movement between the sub-states, the KH-S1-3'ETS^{leuZ}-HFQ sub-assembly is not a perfect rigid body. Therefore, although PCA analysis found the correct modes to describe the continuous movement, the improvement in resolution after rigid body refinement was marginal. A different strategy was used to better resolve the Hfq region of the ternary complex. Two datasets with VPP and two datasets without VPP were collected and each processed separately to get high resolution consensus reconstructions for Bayesian polishing. These consensus reconstructions typically had a well resolved density for the core (3 Å) yet diffuse density for the KH-S1 portal (Figure 10C). Iterative 3D classification and 3D refinement indicated that for VPP-data, the KH-S1-ETS-Hfq region had stronger secondary structure features but limited overall resolution. For the non-VPP data, the overall resolution of the reconstructions was much higher, but it was harder to classify and refine the individual

KH-S1-ETS-Hfq sub-states. These observations are in line with the standard effect of VPPs on data: Low-mid spatial frequency structure factors are boosted due to the phase shift, but high-resolution features are weakened due to the extra carbon layer the electron beam has to travel through. The best results were obtained by exporting and combining polished particle sets from all four datasets (two collected with VPP and two without VPP) in cryoSPARC and running multiple rounds of heterogenous refinement. This strategy gave the best outcome: a series of discrete structures of representing the continuous heterogeneity in the particles, each with well-defined secondary structure features and estimated resolutions going up to 4.4 Å in the KH-S1-3'ETS^{leuZ}-HFQ region for the best resolved sub-state (Figure 10 and Figure 13). It should be noted that all four datasets were collected on the same microscope, with the same magnification and other nominal settings, such as total dose, number of frames per movie, etc. In parallel, polished particles sets for the VPP datasets were combined in cryoSPARC for 3DVA (6 modes, 6Å resolution limit). A detailed overview of the processing pipeline for PNPase-Hfq-3'ETS^{leuZ} is depicted in Figure 11). The PNPase-Hfq-CyaR dataset was processed in Relion 3.0 as described above.

Global resolutions were estimated according to GS-FSC (Gold Standard Fourier shell correlation) between two independently reconstructed half maps ($FSC_{0.143}$). Local resolution variations were estimated with Relion, using the two independent half-maps, or in cryoSPARC (FSC 0.5). For model building, maps were locally filtered according to these locally estimated resolutions. The data processing parameters for apo-PNPase, PNPase-3'ETS^{leuZ} and PNPase-3'ETS^{leuZ}-Hfq are summarized in Table 1.

5.8 Model building and refinement

For initial docking into the apo-PNPase cryo-EM map, MODELLER (Fiser & Sali, 2003) was used to generate a homology model based on the crystal structure of *C. crescentus* PNPase (pdb-ID 4aim) (Hardwick *et al.*, 2012), for which the KH and S1 domains were resolved. Initial docking was performed in UCSF Chimera (Pettersen *et al.*, 2004). A molecular dynamics-based approach, Namdinator (Kidmose *et al.*, 2018), was used to further improve the global fit of the KH-S1 domains. The apo-PNPase structure was then used as a starting model for fitting against the PNPase-sRNA and PNPase-Hfq-3'ETS^{leuZ}/PNPase-Hfq-CyaR cryo-EM maps. For the PNPase sRNA maps, a co-crystal structure of *C.* crescentus PNPase bound to substrate RNA was used as a reference model for the RNA backbone (pdb-ID 4am3) (Hardwick *et al.*, 2010). For the ternary complexes, a crystal structure for *E. coli* Hfq was docked in the cryo-EM maps (pdb-ID 1HK9) and the 3'ETS^{leuZ} RNA backbone was traced in Coot (Emsley *et al.*, 2010). Ramachandran outliers were fixed in Coot, as well as overall geometry. Models were refined with Refmac5 as part of ccpem 1.4.1 (Murshudov *et al*. 2011; Burnley *et al*., 2017) and ISOLDE (Croll, 2018). The quality of the stereochemistry was evaluated via the comprehensive validation tool in Phenix 1.18 (Afonine *et al.*, 2018), which uses molprobity (Chen *et al.*, 2010). Images and movies were made in UCSF Chimera and PyMoL (DeLano, 2012; Pettersen *et al.*, 2004). The statistics for model refinement are displayed in Table 1.

5.9 Gel mobility shift assays (EMSA)

All gel shift assays were carried out at 4°C. RNA was used at 400 nM final concentration and PNPase/Hfq were added at equimolar ratios unless stated otherwise in Figure 20. The binding buffer consisted of 25 mM Tris pH 8, 50 mM NaCl, 20 mM KCl, 100 mM NaCl and 1mM $MgCl_2$. After incubation for 20 minutes at 30°C, native loading

dye (50% (v/v) glycerol, 49.5% binding buffer, 5 mM DTT) was added to the samples before loading them on a native gel. Native electrophoresis was performed in 8% polyacrylamide gels supplemented with 10% glycerol, in Tris-glycine buffer pH 8.5. The gels were run for 4-5 hours at 150V (4°C) and stained for 30-60 min with SYBR Gold (Invitrogen) and visualised with GeneSnap, after excess SYBR Gold was washed from the gel by rinsing with Milli-Q water.

5.10 RNA degradation assays

Degradation assays were performed as described in Bandyra *et al.*, 2016. For the PNPase degradation assays (Figure 15) 200 nM RNA was annealed for 2 minutes at 50°C and mixed with Hfq, if appropriate, in degradation buffer (20 mM Tris, pH 7.5, 100 mM NaCl, 1 mM $MgCl_2$, 1 mM DTT, 2 mM sodium phosphate). Degradation was started by addition of wild type or mutant PNPase at 37 °C. Samples taken from the reaction were quenched at designated time points by adding and equal volume of 0.5 mg/mL proteinase K, dissolved in a buffer containing 100 mM Tris-HCl pH 7.5, 150 mM NaCl, 12.5 mM EDTA and 1% w/v SDS and incubated at 50°C for 20 min. RNA contents corresponding to each timepoint were run on a 8% polyacrylamide, 7.5M urea gel (National Diagnostics) in 1xTBE and visualised using Sybr Gold staining (Invitrogen). Quantification was performed with GeneTools (Syngene).

For the truncated degradosome (RNase E (1-850)/RhlB/Eno) degradation assays the procedure was the same but degradation reactions were prepared in a buffer consisting of 25 mM Tris-HCl pH 7.5, 50 mM NaCl, 50 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, 0.5 U/µL RNase OUT (Invitrogen). RNase III assays were performed according to manufacturer's protocol (ShortCut® RNase III; NEB).

5.11 Immunoprecipitation assays of PNPase

These experiments were carried out by the De Lay group, Department of Microbiology & Molecular Genetics, University of Texas Health Science Center, and are described in Dendooven *et al.,* in preparation.

5.12 *In vivo* RNA stability assays and Northern blot analysis

These experiments were carried out by the De Lay group, Department of Microbiology & Molecular Genetics, University of Texas Health Science Center, and are described in Dendooven *et al.,* in preparation.

6 References

Afonine P.V., Poon B.K., Read R.J., Sobolev O.V., Terwilliger T.C., Urzhumtsev A., & Adams P.D. (2018). Real space refinement in PHENIX for cryo-EM and crystallography. Acta Cryst. *D74*, 531-544.

Afonyushkin, T., Večerek, B., Moll, I., Bläsi, U. & Kaberdin, V.R. (2005) Both RNase E and RNase III control the stability of *sodB* mRNA upon translational inhibition by the small regulatory RNA RyhB. *Nucleic Acids Res.* **33**, 1678–1689.

Andrade, J. M., Pobre, V., Matos, A. M., & Arraiano, C. M. (2012). The crucial role of PNPase in the degradation of small RNAs that are not associated with Hfq. *RNA (New York, N.Y.)*, *18*(4), 844–855. https://doi.org/10.1261/rna.029413.111

Bandyra, K.J., Sinha, D., Syrjanen, J., Luisi, B.F. & De Lay, N.R. (2016) The ribonuclease polynucleotide phosphorylase can interact with small noncoding RNAs in both stabilizing and degradative modes. *RNA* 22, 360-372. PMC4748814

Beckert, B., Turk, M., Czech, A., Berninghausen, O., Beckmann, R., Ignatova, Z., Plitzko, J. M. & Wilson, D. N. (2018). Structure of a hibernating 100S ribosome reveals an inactive conformation of the ribosomal protein S1. *Nature Microbiology*, **3**, 1115–1121. https://doi.org/10.1038/s41564-018-0237-0

Burnley, T., Palmer, C. M., & Winn, M. (2017). Recent developments in the CCP-EM software suite. *Acta crystallographica. Section D, Structural biology*, **73**, 469–477. https://doi.org/10.1107/S2059798317007859

Cameron TA, De Lay NR. (2016) The phosphorolytic exoribonucleases polynucleotide phosphorylase and RNase PH stabilize sRNAs and facilitate regulation of their mRNA targets. *J. Bacteriol.*, 198(24), 3309–3317.

Cameron, T.A., Matz, L.M. and De Lay, N.R. (2018) Polynucleotide phosphorylase: not merely an RNase but a pivotal post-transcriptional regulator. PLos Genetics 14:e1007654.

Cameron, T.A., Matz, L.M. , Sinha, D. and De Lay, N.R. (2019) Polynucleotide phosphorylase promotes the stability and function of Hfq-binding sRNAs by degrading target mRNA-derived fragments. Nucleic Acids Res. – epub.

Chen, H.W., Rainey, R.N., Balatoni, C.E., Dawson, D.W., Troke, J.J., Wasiak, S., Hong, J.S., McBride, H.M., Koehler, C. M., Teitell, M.A. & French, S. W (2006). Mammalian polynucleo- tide phosphorylase is an intermembrane space RNase that maintains mitochondrial homeostasis. Mol Cell Biol. ; *26(22)*:8475–87. https://doi.org/10.1128/MCB.01002-06

Chen, J. & Gottesman, S. (2017) Hfq links translation repression to stress-induced mutagenesis in *E. coli*. Genes Dev.. https://doi.org/10.1101/gad.302547.117 PMID: 28794186; PubMed Central PMCID: PMCPMC5580658

Chen, J., Noble, A.J., Kang, J.Y. and Darst, S. (2019) Eliminating effects of particle adsorption to the air/water interface in single-particle cryo-electron microscopy: bacterial RNA polymerase and CHAPSO. Journal of Structural Biology: X, Volume1, 100005, https://doi.org/10.1016/j.yjsbx.2019.100005.

Chen VB*, et al.* (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66:12-21.

Chen X, Taylor DW, Fowler CC, Galan JE, Wang HW, Wolin SL (2013). An RNA degradation machine sculpted by Ro autoantigen and noncoding RNA. Cell **153**:166–177. doi:10.1016/j.cell.2013.02.037.

Cheng *et al.* (2018) Mitochondrial trafficking and processing of telomerase RNA TERC. *Cell Reports* 24, 2589-2595.

Cheng, Z.F. and Deutscher, M.P. (2003) Quality control of ribosomal RNA mediated by polynucleotide phosphorylase and RNase R. Proc. Natl. Acad. Sci. U.S.A. 100, 6388-6393.

Chou-Zheng, L. and Hatoum-Aslan, A. (2019) A type III-A CRISPR-Cas system employ degradosome nucleases to ensure robust immunity. eLife 8 pii: e45393.

Chujo, T., Ohira, T., Sakaguchi, Y., Goshima, N., Nomura, N., Nagao, A. and Suzuki, T. (2012) LRPPRC/SLIRP suppresses PNPase-mediated mRNA decay and promotes polyadenylation in human mitochondria. Nucleic Acids Res. 40, 8033-8047.

Croll. T.I. (2018) ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr D Struct Biol* 74:519-530.

Danchin, A. (1997) Comparison between the Escherichia coli and Bacillus subtilis genomes suggests that a major function of polynucleotide phosphorylase is to synthesize CDP. (PMID:9179491)

Danev, R., & Baumeister, W. (2016). Cryo-EM single particle analysis with the Volta phase plate. *eLife*, *5*, e13046. https://doi.org/10.7554/eLife.13046

Dar, D. and Sorek, R. (2018) High-resolution RNA 3'-ends mapping of bacterial Rho-dependent transcripts. *Nucleic Acids Res.* 46, 6797-6805.

DeLano WL (2002) The PyMOL molecular graphics system. *http://wwwpymolorg*.

De Lay, N. & Gottesman, S. (2009) The Crp-Activated Small Noncoding Regulatory RNA CyaR (RyeE) Links Nutritional Status to Group Behavior. Journal of Bacteriology; *191 (2):* 461-476; DOI**:** 10.1128/JB.01157-08

De Lay N, Gottesman S. (2011) Role of polynucleotide phosphorylase in sRNA function in *Escherichia coli*. *RNA*, 17(6), 1172–1189.

Dendooven, T., Luisi, B.F. (2017) RNA search engines empower the bacterial intranet. Biochem Soc Trans; *45 (4)*: 987–997. doi: https://doi.org/10.1042/BST20160373

Deutscher, M.P. (2015) How bacterial cells keep ribonucleases under control. FEMS Microbiol. Rev. *012*: 350–361.

Dimastrogiovanni, D., Frohlich, K.S., Bandyra, K.J., Bruce, H.A, Hohensee, S., Vogel, J. and Luisi, B.F. (2014) Recognition of the small regulatory RNA RydC by the bacterial Hfq protein. *eLife* doi 10.7554/eLife.05375. PMID 25551292. PMC4337610

D'Imprima, E., Floris, D., Joppe, M., Sánchez, R., Grininger, M., & Kühlbrandt, W. (2019). Protein denaturation at the air-water interface and how to prevent it. *eLife*, *8*, e42747. https://doi.org/10.7554/eLife.42747

Dhir, A. *et al.* (2018) Mitochondrial double-stranded RNA triggers antiviral signaling in humans. *Nature* 560, 238-242.

Duval, M., Korepanov, A., Fuchsbauer, O., Fechter, P., Haller, A., Fabbretti, A., Choulier, L., Micura, R., Klaholz, B.P., Romby, P., Springer, M. and Marzi, S. (2013) Escherichia coli ribosomal protein S1 unfolds structured mRNAs onot the ribosome for active translation initiation. PloS Biol. E1001731.

Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M. Y., … Sali, A. (2006). Comparative protein structure modeling using Modeller. Current protocols in bioinformatics, Chapter 5, Unit–5.6. doi:10.1002/0471250953.bi0506s15

Fazal, F.M., Koslover, D.J., Luisi, B.F., Block, S.M. (2015) Direct observation of processive exoribonuclease motion using optical tweezers. *Proc. Natl. Acad. Sci. U.S.A.* 112, 15101-15106. PMID 26598710. PMC4679025.

Figueroa-Bossi, N., Valentini, M., Malleret, L., Fiorini, F., & Bossi, L. (2009). Caught at its own game: regulatory small RNA inactivated by an inducible transcript mimicking its target. *Genes & development*, *23*(17), 2004–2015. https://doi.org/10.1101/gad.541609

Fiser, A. & Šali, A. B. T.-M. in E. Modeller: Generation and Refinement of Homology-Based Protein Structure Models. in Macromolecular Crystallography, Part D *374*, 461–491

Gatewood ML, and Jones GH (2010). (p) ppGpp inhibits polynucleotide phosphorylase from streptomyces but not from Escherichia coli and increases the stability of bulk mRNA in Streptomyces coelicolor. J Bacteriol **192**:4275–4280. doi:10.1128/JB.00367-10.

Golzarroshan, B., Lin, C. L., Li, C. L., Yang, W. Z., Chu, L. Y., Agrawal, S., & Yuan, H. S. (2018). Crystal structure of dimeric human PNPase reveals why disease-linked mutants suffer from low RNA import and degradation activities. *Nucleic acids research*, *46*(16), 8630–8640. https://doi.org/10.1093/nar/gky642

Grunberg-Manago, M. & Ochoa, S. (1995). Enzymatic synthesis and breakdown of polynucleotides; polynucleotide phosphorylase. J Am Chem Soc.; *77*:3165–6.

Hardwick, S., Gubbey, T., Hug, I., Jenal, U. & Luisi, B.F. (2012) Crystal structure of *Caulobacter crescentus* polynucleotide phosphorylase reveals a mechanism of RNA substrate channelling and RNA degradosome assembly. *Open Biology* **2**, 120028. doi:10.1098/rsob.120028. PMC3376730

Holmqvist, E., Reimegard, J., Sterk, M., Grantcharova, N., Romling, U. & Wagner, E.G. (2010). Two antisense RNAs target the transcriptional regulator CsgD to inhibit curli synthesis. EMBO J; *29(11*):1840–50. https://doi.org/10.1038/emboj.2010.73 PMID: 20407422; PubMed Central PMCID: PMCPMC2885931.

Jagodnik, J., Chiaruttini, C. Guillier, M. (2017). Stem-Loop Structures within mRNA Coding Sequences Activate Translation Initiation and Mediate Control by Small Regulatory RNAs. Mol Cell 68(1): 158-170 e153.

Jin, Y., Watt, R. M., Danchin, A., & Huang, J. D. (2009). Small noncoding RNA GcvB is a novel regulator of acid resistance in Escherichia coli. BMC genomics, 10, 165. https://doi.org/10.1186/1471-2164-10-165

Johansen, J., Eriksen, M., Kallipolitis, B. & Valentin-Hansen, P. (2008) Down-regulation of outer membrane proteins by noncoding RNAs: unraveling the cAMP-CRP- and sigmaE-dependent CyaR-ompX regulatory case. J Mol Biol. ;*383(1)*:1-9. doi:10.1016/j.jmb.2008.06.058

Jorgensen, M.G., Thomason, M.K., Havelund, J., Valentin-Hansen, P. & Storz, G. (2013) Dual function of the McaS small RNA in controlling biofilm formation. Genes Dev. 2013; *27(10)*:1132–45. https://doi.org/10.1101/gad.214734.113 PMID: 23666921; PubMed Central PMCID: PMCPMC3672647.

Kastner, B., Fischer, N., Golas, M. *et al.* GraFix: sample preparation for single-particle electron cryomicroscopy. *Nat Methods* **5,** 53–55 (2008). https://doi.org/10.1038/nmeth1139.

Kidmose RT, Juhl J, Nissen P, Boesen T, Karlsen JL, Pedersen BP (2019) Namdinator - Automatic Molecular Dynamics flexible fitting of structural models into cryo-EM and crystallography experimental maps. IUCrJ 6, 526-531.

Kim, W. and Lee, Y. (2020) Mechanism for coordinate regulation of *rpoS* by sRNA-sRNA interaction in *Escherichia coli*. *RNA Biology* **17**, 176-187.

Lalaouna, D., Carrier, M.C., Semsey, S., Brouard, J.S., Wang, J., Wade, J.T. and Masse, E. (2015) A 3' external transcribed space in a tRNA transcript acts as a sponge for small RNAs to prevent transcriptional noise. Mol. Cell 58, 393-405.

Lee, G., Bratkowski, M. A., Ding, F., Ke, A. & Ha, T. (2012) Elastic coupling between RNA degradation and unwinding by an exoribonuclease. Science*. 336*, 1726–1729.

Lengyel, P., Speyer, J.F. & Ochoa, S. (1961) Synthetic polynucleotides and the amino acid code. Proc. Natl. Acad. Sci. U. S. A. 47: 1936–1942.

Masse E., Escorcia F.E., Gottesman S. (2003) Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*. *Genes Dev*. **17**, 2374–2383.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., & Vagin, A. A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta crystallographica. Section D, Biological crystallography*, **67**, 355–367. https://doi.org/10.1107/S0907444911001314

Nishino, K., Yamasaki, S., Hayashi-Nishino, M. & Yamaguchi, A. (2011) Effect of overexpression of small non-coding DsrA RNA on multidrug efflux in *Escherichia coli*. J Antimicrob Chemother; *66(2)*:291–6. https://doi.org/10.1093/jac/dkq420 PMID: 21088020.

Nurmohamed, S., Vaidialingam, B., Callaghan, A.J., and Luisi, B.F. (2009) Crystal structure of *Escherichia coli* polynucleotide phosphorylase core bound to RNase E, RNA and manganese: implications for catalytic mechanism and RNA degradosome assembly. *J. Mol. Biol*. **389**, 18-33.

Nurmohamed, S., Vincent, H.A., Titman, C.M., Chandran, V., Pears, M.R., Du, D., Griffin, J.L., Callaghan, A.J. and Luisi, B.F. (2011). Polynucleotide phosphorylase activity may be modulated by metabolites in *Escherichia coli*. *J. Biol. Chem*. **286**, 14315-14323.

Pettersen EF*, et al.* (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25:1605-1612.

Plocinski *et al.* (2019) Proteomic and transcriptomic experiments reveal an essential role of RNA degradosome complexes in shaping the transcriptome of Mycobacterium tuberculosis. Nucleic Acid Res. 47, 5892-5905.

Punjani, A., Rubinstein, J. L., Fleet, D. J., & Brubaker, M. A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature methods*, **14**, 290–296. https://doi.org/10.1038/nmeth.4169

Punjani, A. & Fleet, D.J. (2020). 3D Variability Analysis: Directly resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM images. *bioRxiv*. doi:10.1101/2020.04.08.032466

Razinkov, I., Dandey, V., Wei, H., Zhang, Z., Melnekoff, D., Rice, W. J., Wigge, C., Potter, C. S., & Carragher, B. (2016). A new method for vitrifying samples for cryo-EM. Journal of structural biology, 195(2), 190–198. https://doi.org/10.1016/j.jsb.2016.06.001

Santiago-Frangos, A., Jeliazkov, J.R., Gray, J.J., and Woodson, S.A. (2017). Acidic C-terminal domains autoregulate the RNA chaperone Hfq. Elife *6*, e27049.

Santiago-Frangos, A., Kavita, K., Schu, D. J., Gottesman, S., and Woodson, S. A. (2016). C-terminal domain of the RNA chaperone Hfq drives sRNA competition and release of target RNA. Proc. Natl. Acad. Sci. U.S.A. *113*, E6089–E6096.

Santiago-Frangos, A., Fröhlich, K.S., Jeliazkov, J.R., Małecka, E.M., Marino, G., Gray, J.J., Luisi, B.F., Woodson, S.A., Hardwick, S.W. (2019) Caulobacter crescentus Hfq structure reveals a conserved mechanism of RNA annealing regulation. *Proc Natl Acad Sci*. *116*:10978–10987.

Sato, R., Arai-Ichinoi, N., Kikuchi, A., Matsuhashi, T., Numata-Uematsu, Y., Uematsu, M., Fujii, Y., Murayama, K., Ohtake, A., Abe, T., & Kure, S. (2018). Novel biallelic mutations in the PNPT1 gene encoding a mitochondrial-RNA-import protein PNPase cause delayed myelination. *Clinical genetics*, *93*(2), 242–247. https://doi.org/10.1111/cge.13068

Schu, D.J., Zhang, A., Gottesman, S. and Storz, G. (2015) Alternative Hfq-sRNA interaction modes dictate alternative mRNA recognition. *EMBO J.* **34**, 2557-2573.

Schulz, E. C., & Barabas, O. (2014). Structure of an Escherichia coli Hfq:RNA complex at 0.97 Å resolution. *Acta crystallographica. Section F, Structural biology communications*, *70*(Pt 11), 1492–1497. https://doi.org/10.1107/S2053230X14020044

Shi, Z., Yang, W. Z., Lin-Chao, S., Chak, K. F., & Yuan, H. S. (2008). Crystal structure of Escherichia coli PNPase: central channel residues are involved in processive RNA degradation. *RNA 14*(11), 2361–2371. https://doi.org/10.1261/rna.1244308

Shimada, E., Ahsan, F. M., Nili, M., Huang, D., Atamdede, S., TeSlaa, T., Case, D., Yu, X., Gregory, B. D., Perrin, B. J., Koehler, C. M., & Teitell, M. A. (2018). PNPase knockout results in mtDNA loss and an altered metabolic gene expression program. *PloS one*, *13*(7), e0200925. https://doi.org/10.1371/journal.pone.0200925.

Shokolenko, I.N. & Alexeyev, M.F. (2017) Mitochondrial transcription in mammalian cells. Front Biosci; *2*2:835–53. PMID: 27814650; PubMed Central PMCID: PMCPMC5653244.

Sesto, N., Touchon, M., Andrade, J.M., Kondo, J., Rocha, E.P., Arraiano, C.M., Archambaud, C., Westhof, E., Romby, P. and Cossart, P. (2014) A PNPase dependent CRISPR system in Listeria. *PLoS Genet*. 10: e1004065.

Silva, S., Camino, L.P. and Aguilera, A. (2018) Human mitochondrial degradosome prevents harmful mitochondrial R loops and mitochondrial genome instability. *Proc Natl Acad Sci. U.S.A.* 115, 11024-11029.

Sim, S. and Wolin, S.L. (2018) Bacterial Y RNAs: Gates, Tethers, and tRNA mimics. Microbiol. Spectr. 6 RWR-0023-2018.

Söderbom, F. & Wagner, E. G. H. Degradation pathway of CopA, the antisense RNA that controls replication of plasmid R1. *Microbiology* **144**, 1907–1917 (1998).

Spickler C, Mackie GA (2000) Action of RNase II and polynucleotide phosphorylase against RNAs containing stem-loops of defined structure. *Journal of bacteriology* 182 : 2422–7

Stickney LM, Hankins JS, Miao X, Mackie GA. (2005) Function of the conserved S1 and KH domains in polynucleotide phosphorylase. J Bacteriol 187:7214 –7221.

Stone, C.M., Butt, L.E., Bufton, J.C., Lourenco, D.C., Gowers, D.M., Pickford, A.R., Cox, P.A., Vincent, H.A., and Callghan, A.J. (2017) Inhibition of homologous phosphorolytic ribonucleases by citrate may represent an evolutionarily conserved communicative link between RNA degradation and central metabolism. *Nucleic Acids Res.* 45, 4644-4666.

Symmons, M.F., Jones, G.H. and Luisi, B.F. (2000). A duplicated fold is the structural basis for polynucleotide phosphorylase catalytic activity, processivity, and regulation. *Structure* **8**, 1215-1226.

Symmons, M.F., Williams, M.G., Luisi, B.F., Jones, G.H., and Carpousis, A.J. (2002) Running rings around RNA: a superfamily of phosphate-dependent RNases. *Trends Biochem. Sciences* **27**, 11-18.

Tan, Y. Z., Baldwin, P. R., Davis, J. H., Williamson, J. R., Potter, C. S., Carragher, B., & Lyumkis, D. (2017). Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nature methods*, *14*(8), 793–796. doi:10.1038/nmeth.4347

Thomason, M.K., Fontaine, F., De Lay, N. & Storz, G. (2012) A small RNA that regulates motility and biofilm formation in response to changes in nutrient availability in *Escherichia coli*. Mol Microbiol.; *84(1)*:17– 35. https://doi.org/10.1111/j.1365-2958.2012.07965.x PMID: 22289118; PubMed Central PMCID: PMCPMC3312966.

Tuckerman JR, Gonzalez G, Gilles-Gonzalez MA (2011) Cyclic di-GMP activation of polynucleotide phosphorylase signal-dependent RNA processing. J Mol Biol **407**:633–639. doi:10.1016/j.jmb.2011.02.019.

Urbanowski, M.L., Stauffer, L.T. and Stauffer, G.V. (2000), The gcvB gene encodes a small untranslated RNA involved in expression of the dipeptide and oligopeptide transport systems in Escherichia coli. Molecular Microbiology, 37: 856-868. doi:10.1046/j.1365-2958.2000.02051.x

Viegas, S. C., Pfeifer, V., Sittka, A., Vogel, J. & Arraiano, C.M. (2007) Characterization of the role of ribonucleases in Salmonella small RNA decay . *Nucleic Acids Res.* **35**, 7651–7664.

Vedrenne, V., Gowher, A., De Lonlay, P., Nitschke, P., Serre, V., Boddaert, N., Altuzarra, C., Mager-Heckel, A. M., Chretien, F., Entelis, N., Munnich, A., Tarassov, I., & Rötig, A. (2012). Mutation in PNPT1, which encodes a polyribonucleotide nucleotidyltransferase, impairs RNA import into mitochondria and causes respiratory-chain deficiency. *American journal of human genetics*, *91*(5), 912–918. https://doi.org/10.1016/j.ajhg.2012.09.001

Vogel, J., Luisi, B. Hfq and its constellation of RNA. *Nat Rev Microbiol* **9,** 578–589 (2011). https://doi.org/10.1038/nrmicro2615

von Ameln, S., Wang, G., Boulouiz, R., Rutherford, M. A., Smith, G. M., Li, Y., Pogoda, H. M., Nürnberg, G., Stiller, B., Volk, A. E., Borck, G., Hong, J. S., Goodyear, R. J., Abidi, O., Nürnberg, P., Hofmann, K., Richardson, G. P., Hammerschmidt, M., Moser, T., Wollnik, B., … Kubisch, C. (2012). A mutation in PNPT1, encoding mitochondrial-RNA-import protein PNPase, causes hereditary hearing loss. *American journal of human genetics*, *91*(5), 919–927. https://doi.org/10.1016/j.ajhg.2012.09.002

Wang, G., Chen, H. W., Oktay, Y., Zhang, J., Allen, E. L., Smith, G. M., Fan, K. C., Hong, J. S., French, S. W., McCaffery, J. M., Lightowlers, R. N., Morse, H. C., 3rd, Koehler, C. M., & Teitell, M. A. (2010). PNPASE regulates RNA import into mitochondria. *Cell*, *142*(3), 456–467. https://doi.org/10.1016/j.cell.2010.06.035

Wang, M., Herrmann, C.J., Simonovic, M., Szklarczyk, D. & von Mering, C. (2015) Version 4.0 of PaxDb: Protein abundance data, integrated across model organisms, tissues, and cell-lines. Proteomics *15*: 3163–3168

Waters, L.S. & Storz, G. (2009) Regulatory RNAs in Bacteria. Cell 136: 615–628

Webb, B., Sali, A. (2016) Comparative Protein Structure Modeling Using MODELLER, In Current Protocols in Bioinformatics , volume 54, Hoboken, NJ, USA: John Wiley & Sons, Inc., pp. 5.6.1–5.6.37

Zhang, A., Schu, D.J., Tjaden, B.C., Storz, G. and Gottesman, S. (2013) Mutations in interaction surfaces differentially impact E. coli Hfq association with small RNAs and their mRNA targets. *J. Mol. Biol.* **425**, 3678-3697.

Zivanov, J., Nakane, T., Forsberg, B., Kimanius, D., Hagen, W.J.H., Lindahl, E. and Scheres, S.H.W. (2018) RELION-3: new tools for automated high-resolution cryo-EM structure determination. *eLife* **7**:e42166.

Zuker M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic acids research*, *31*(13), 3406–3415. https://doi.org/10.1093/nar/gkg595

Chapter III: Un-structural Biology of the bacterial RNA degradosome

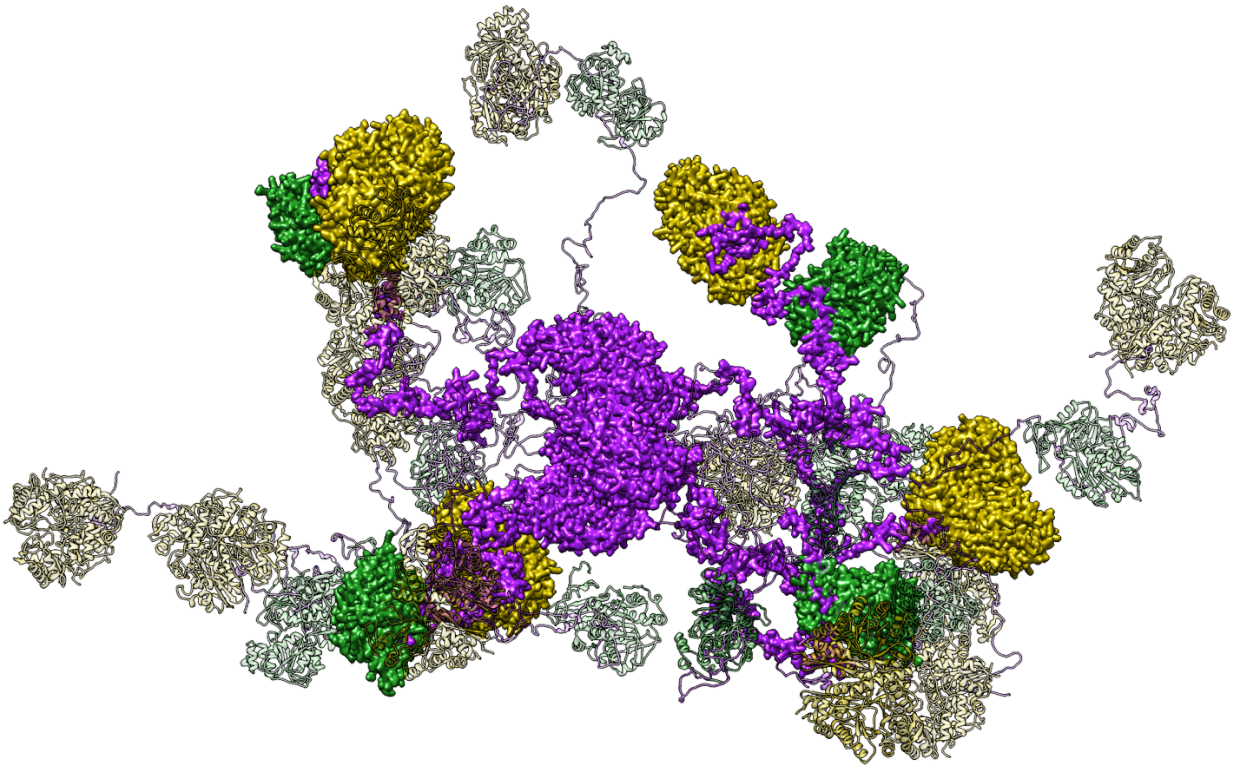


Table of Contents

1 Abstract

In organisms from all domains of life, multi-enzyme assemblies play central roles in defining the lifetime of transcripts and in facilitating RNA-mediated control of gene regulation. Bacteria of extremely diverse lineages have a key regulatory assembly known as the RNA degradosome, the scaffold of which is often provided by the endoribonuclease RNase E. This enzyme contributes a natively unstructured region which by mass, corresponds to about a fifth of the assembly. This unstructured property is strongly sustained in a billion years of divergence, in marked contrast to the lack of sequence conservation, and is likely to be beneficial at different evolutionary timescales. Consequently, the unstructured character makes the RNA degradosome unamenable to conventional structural approaches. In this chapter, integrative solution scattering approaches were developed to map the ensemble behaviour of the RNA degradosome. In addition, single particle analysis of the catalytic core of the RNA degradosome reveals its conformational landscape. Lastly Cryo-electron tomography-based strategies were explored to set a foundation for future structural studies of the membrane bound RNA degradosome.

2 Introduction

The central dogma in Biology states that a gene is transcribed to mRNA, which is then translated to a gene product. The fate of the mRNA and how it affects expression of genes, i.e. post-transcriptional gene regulation, has been shown to be just as important as the acts of transcription and translation themselves (Felden & Paillard, 2017). Coordination of RNA lifetime, together with adjustments to rates of synthesis, maintains transcript homeostasis, enables rapid change of gene expression, and signals cellular status (Adler & Alon, 2018; Sun *et al.,* 2012; Haimovich *et al.*, 2013; Peréz-Ortìn *et al.*, 2007; Schmidt & Jensen, 2018; Tudek *et al.*, 2018; Bresson & Tollervey, 2018). Some transcripts, like rRNAs and tRNAs, are relatively stable. mRNA half-lives, on the other hand, are strongly dependent on growth phase and environmental factors (Dressaire *et al.*, 2013). Beyond the codon pattern that mRNAs carry for protein translation, they can also encode structural features that define lifetime or the potential to be translated before turnover, and its maturation can unmask hidden information made available through interactions with other transcripts (Chao *et al.* 2017; Dendooven *et al.* 2017). The maturation and turnover of this diverse pool of transcripts is tightly regulated in cells from all domains of life (Dendooven *et al.*, 2020; Tejada-Arranz *et al.*, 2020).

Prokaryotic RNA processing has been extensively studied using bacterial model organisms such as the Gram-negative *Escherichia coli* and the Gram-positive *Bacillus subtilis*, which arose from highly divergent bacterial lineages. RNA decay in these bacteria has become centred around multi-enzyme complexes over the course of evolution. These assemblies are often built on a membrane-bound multi-domain ribonuclease scaffold: RNase E in the case of most Gram-negative bacteria (Bandyra *et al.*, 2013) and RNase Y in Gram-positive microorganisms such as *Bacillus subtilis* and the pathogenic species *Staphylococcus aureus* (Lehnik-Habrink *et al.* 2011; Commichau *et al.*, 2009; Durand *et al.*, 2015; Roux *et al.*, 2011). RNase E and RNase Y are functionally analogous hydrolytic endonucleases with a distinct preference for 5'-monophosphorylated, single stranded RNA species (Hui *et al.*, 2014; Mackie, 2012; Shahbabian *et al.* 2009*)*. The RNase E and RNase Y ribonucleases both assemble into

functionally conserved multi-enzyme machineries, often referred to as RNA degradosomes. The widespread occurrence of these RNA degrading assemblies among bacteria highlights their functional importance and the composition and organisation of bacterial RNA degradosomes has changed markedly during evolution. Even within a given organism, the degradosome composition can vary with growth phase and environmental conditions (Aït–Bara & Carpousis, 2015). Interestingly, bacterial RNA degradosomes have both functional and structural similarities to the eukaryotic exosome and the RISC (RNA-Induced Silencing Complex) machinery (Bandyra *et al.* 2018, Dendooven *et al.*, 2020). Given that the degradosome is an essential bacterial machine, it is evident that its activity is under tight control. Small regulatory RNAs (sRNAs) and RNA chaperones have been shown to coregulate RNA degradation by complementing canonical RNA degradosome assemblies (Dendooven & Luisi, 2017). Moreover, negative feedback loops, bacterial inhibitors and even phage encoded effector molecules are known to affect degradosome activity (Dendooven & Lavigne, 2019; Dendooven *et al. 2017;* Van den Bossche *et al.*, 2016).

In this chapter, approaches to explore the structural layout of the *E. coli* RNA degradosome will be presented. The *E. coli* RNA degradosome was first discovered by Carpousis *et al.* in 1994, who found that RNase E forms the core unit of the assembly (Figure 1A and B). RNase E has a structured catalytic domain, located in the N-terminal half, which oligomerises in functional tetramers, and performs the initial cleavage of RNA substrates (Callaghan *et al.*, 2005). The other portion of RNase E, the C-terminal half, serves as a scaffold for the partner enzymes of the degradosome and is predicted to be natively unstructured (Bandyra *et al.*, 2013; Bruce *et al.*, 2018). One partner enzyme of the RNA degradosome is the DEAD-box helicase RhlB, which is a canonical degradosome component (Figure 1B) (Py *et al.*, 1996). In fact, over the years it has become clear that RNA degradosome assemblies, both in Gram-negative and Gram-positive bacteria, consists of at least a ribonuclease and a helicase of the DEAD-box family (Aït-Bara & Carpousis, 2015). The latter helps unfolding secondary structures in RNA substrates and may help to remodel ribonucleoprotein complexes, making cleavage sites accessible for the ribonuclease. Often bacterial degradosomes have more than one ribonuclease, each with different but complementary enzymatic activities. PNPase, which has been extensively studied in Chapter II, is the second degradosome ribonuclease in *E. coli* and is also tethered to the RNase E scaffold domain. As described earlier, PNPase is an exoribonuclease that catalyses the 3 ’->5 ’ phosphorolytic cleavage of substrate RNAs (Figure 1B). Finally, several degradosomes also carry a metabolic enzyme, such as aconitase (*Caulobacter crescentus*) or phosphofructokinase (*B. subtilis* and *S. aureus*) (Voss *et al.* 2014, Lehnik-Habrink *et al.*, 2011; Roux *et al.*, 2011). In *E. coli*, Enolase fulfils the role of metabolic enzyme in the degradosome (Figure 1B), although its role in RNA turnover is still elusive. Studies have suggested that Enolase is involved in coupling the energy status of the cell to RNA degradation (Morita *et al.*, 2004; Bruce *et al.*, 2018). In addition, association of Enolase with RNase E is required for cell division control under anaerobic conditions (Murashko and Lin-Chao, 2017). Together, these four enzymes (RNase E, RhlB, PNPase and Enolase) constitute the canonical *E. coli* RNA degradosome, a complex molecular machine and main actor in steady state turnover of the cellular mRNA pool and post-transcriptional regulation of gene expression.

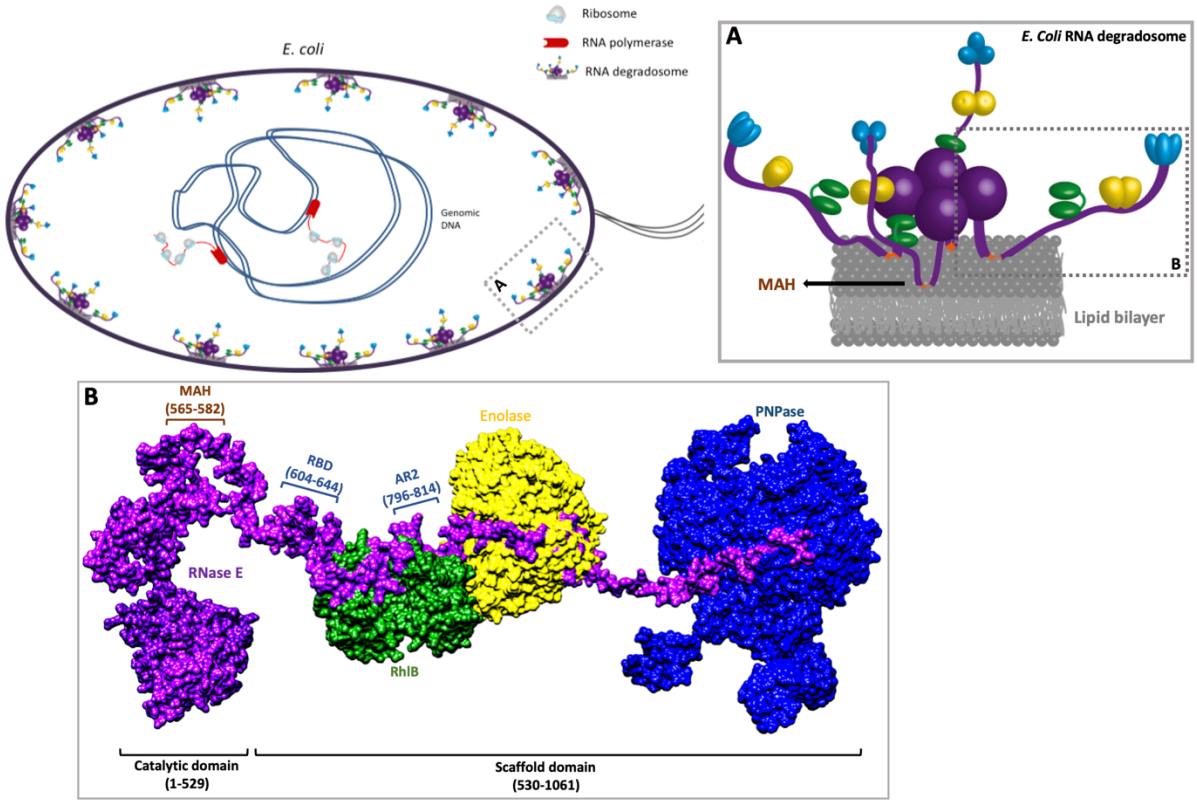


Figure 1: ***Schematic of membrane association and architecture of the E. coli RNA degradosome. (A)*** *The degradosome is a homo-tetrameric multi-enzyme assembly that is tethered to the bacterial inner membrane via the RNase E amphipathic Membrane Attachment Helix (MAH).* ***(B)*** *The core component of the degradosome is the endoribonuclease RNase E. RNase E consists of an N-terminal catalytic domain, responsible for initial RNA substrate cleavage, and a C-terminal scaffold domain, which is predicted to be natively unstructured. The scaffold domain is a mosaic of microdomains involved in recruiting other degradosome components and RNA substrates. The Membrane Attachment Helix (MAH) embeds the degradosome into the lipid bilayer of the inner membrane. Two RNA binding sites (RBD and AR2) capture RNA substrates for degradation. Interaction sites for the other components of the canonical RNA degradosome, the DEAD-box helicase RhlB, the glycolytic enzyme Enolase and the exoribonuclease PNPase complete the RNase E scaffold domain. Figure adapted from Bandyra et al., 2013; Dendooven & Luisi, 2017; Dendooven & Lavigne, 2019.*

Although this chapter will mainly address large sub-assemblies of the RNA degradosome, there will be a focus on strategies to study the C-terminal scaffold domain of RNase E. The RNase E C-terminal half has gained substantial interest over the last years, mostly due to its peculiar features beyond being a scaffold region for degradosome components. On top of recognition sites for RhlB, Enolase and PNPase, it contains two RNA binding sites, AR2 and RBD, that capture RNA substrates for processing (Garrey *et al.*, 2009; Leroy *et al.*, 2002) (Figure 1B). It has been suggested that these RNA binding sites can cooperate with RhlB in the degradosome to assist in substrate capture and unwinding as well (Khemici & Carpousis, 2004). The RNase E scaffold domain also contains a short amphipathic α-helical domain that interacts with the *E. coli* inner membrane phospholipids, serving as a membrane anchor for the degradosome (Figure 1A and B). Membrane localization of the RNA degradosome has been shown to contribute additional regulatory layers to RNA turnover in the cell and will be further discussed below (section 2.1). In between these microdomains, the RNase E C-terminal half is rich in short linear motifs (SLiMs) that are inherent to intrinsically disordered regions (IDRs) in proteins. Notably, there are clear evolutionary drives for RNase E to

sustain its long, natively disordered scaffold domain (Aït-Bara *et al.*, 2015). What constitutes these evolutionary drives will also be discussed below (section 2.2). Lastly the RNase E scaffold domain was found to be involved in the formation of putative RNA surveillance assemblies (Tsai *et al.*, 2012), and methods to reconstitute these *in vitro* are presented in this chapter. These RNA degradosome based surveillance assemblies, which will be introduced in section 2.3, are believed to prevent spurious transcripts from being transcribed or translated.

2.1 Sub-cellular localization of the RNA degradosome and implications for gene regulation

It seems likely that RNA stability could be controlled, in part, by its cellular location and concomitant proximity to the RNA decay machinery. In many γ-proteobacteria, such as *E. coli*, the degradosome was shown or predicted to associate with the inner membrane (Aït-Bara *et al.*, 2015) (Figure 1). In the Gram-positive bacteria *B. subtilis* and *S. aureus*, RNase Y based degradosomes are also predicted to associate with the cell membrane (Hunt *et al.*, 2006; Aït-Bara *et al.*, 2015). In the α-proteobacterium *C. crescentus*, however, RNase E lacks an amphipathic helix required for membrane-association and was found to be cytosolic (Bayas *et al.*, 2018). Interestingly, the *E. coli* degradosome is highly mobile on the membrane surface but forms transient, punctuate loci that are likely to be centres of RNA turnover (Strahl *et al.*, 2015). These transient degradation clusters were postulated to be cooperative degradation bodies that form during substrate turnover and diffuse upon depletion of the target RNA.

Preliminary proof for a regulatory role for membrane association of the degradosome was obtained by disabling the RNase E membrane attachment helix in *E. coli* (Hadjeras *et al.*, 2019). It was clear that this engineered cytoplasmic form of the degradosome severely disrupts RNA turnover, slowing it down globally, yet increasing turnover of exposed RNAs, i.e. RNAs that are devoid of bound ribosomes (Hadjeras *et al.*, 2019). Light microscopy studies further showed that transcripts encoding inner membrane proteins had an increased stability in a cytoplasmic degradosome background (Moffitt *et al.*, 2016). In a normal bacterial cell these RNAs migrate co-translationally to the inner membrane via signal peptides, where they are more exposed to the membrane-bound degradosome, resulting in shorter half-lives (Moffit *et al.*, 2016). In summary, these studies suggest that membrane association of the degradosome adds a spatial layer to post-transcriptional gene regulation, resulting in an increased turnover of RNA substrates that reside in proximity to the bacterial inner membrane.

2.2 Evolutionary forces that keep degradosomes unstructured and opportunities for spontaneous compartmentalisation

Perhaps the most peculiar feature of the bacterial RNA degradosome is the conserved natively unstructured character of the RNase E scaffold domain. In the literature, two hypotheses have been advanced to explain the evolutionary drive that favours an unstructured scaffold domain. The first hypothesis starts from the observation that the RNA degradosome, albeit functionally conserved, has a wide range of canonical partner enzymes over different Gram-negative and Gram-positive bacterial phylae (Aït-Bara *et al.*, 2015). It is clear that the degradosome has continuously switched components over the course of evolution. For this to have occurred, the microdomains

that punctuate the RNase E scaffold domains would have to evolve to attract new degradosome partners and discard others. It is likely that the unstructured character of the scaffold domain reduces evolutionary constraints on doing so because there is less restrictive pressure for the evolving microdomains to fold into a globally stable and functional structure (Brown *et al.*, 2002; De Pristo *et al.*, 2005; Marcaida *et al.*, 2006). The consequences of mutations in and around these microdomains are acted upon by natural selection without the thermodynamic costs intrinsic to folded proteins. In other words, the unstructured RNase E scaffold domain can evolve faster than, for example, the structured catalytic half of RNase E. The sequence divergence of the C-terminal scaffold domain as opposed to the N-terminal catalytic domain of RNase E supports this hypothesis.

The second hypothesis is based on the recent finding that membrane-less compartmentalisation of regulatory processes can occur in cells (Banani *et al.*, 2017). In eukaryotes, organelles compartmentalise and concentrate metabolic processes to increase their efficiency and prevent unwanted reactions. In bacteria, physically separated compartments do not occur, with a few exceptions (magnetosomes, spores, etc.) (Cornejo *et al.*, 2014). However, a secondary, more subtle mechanism of subcellular compartmentalisation has recently been described in eukaryotes and prokaryotes (Banani *et al.*, 2017). RNA-binding proteins can cluster in so called membrane-less organelles upon interaction with their RNA targets, and this happens via liquid-liquid phase separation (LLPS) (Lin *et al.*, 2015; Boeynaems *et al.*, 2018). In eukaryotes for example, large cytoplasmic granules can form under stress conditions (Guzykowski *et al.*, 2019), often referred to as messenger ribonucleoprotein (mRNP) granules. It has been shown that many RNA binding proteins, RNases, helicases and substrate RNAs can initiate LLPS to form a sequestered processive mega body in the cell. Concentrated processes in these mRNP granules seal the fate of substrate RNAs, whether it is turnover or storage (Guzykowski *et al.*, 2019). In addition, it was found that within LLPS organelles the environment can be altered to behave more like an organic solvent rather than water, altering the stability of substrate RNA secondary structures (Nott *et al.*, 2016). It is not yet fully understood what the physico-chemical prerequisites are for RNA-containing LLPS to occur. However, two common denominators in the LLPS literature are proteins with Intrinsically disordered regions and RNA binding domains (Protter *et al.*, 2018, Banani *et al.*, 2017). For example, solid state NMR studies of the RNA binding protein FUS (Fused in Sarcoma), which forms liquid-liquid phase separated bodies *in vivo*, support a model in which LLPS is mediated by the flexible extensions from its structured core (Murray *et al.*, 2017; Murthy *et al.*, 2019).

Recent proof of bacterial liquid-liquid phase separated RNA degradosome bodies was published for the aquatic Gram-negative bacterium *Caulobacter crescentus* (Al-Husini *et al.*, 2018). It was found that formation of these bodies is dependent on RNA-binding and can be reversed by RNA turnover. As stated above, RNA degradosomes have a conserved unstructured character in most bacterial lineages, containing many RNA-binding sites, and therefore are good candidate mediators of LLPS within a bacterial cell. A similar phenomenon has been observed for the membrane associated *E. coli* RNA degradosome (Strahl *et al.*, 2015), where degradosomes form transient clusters during RNA turnover, as explained above. For the latter, the liquid-liquid phase separated degradosome 'nano-organelles' are likely to be extended in two dimensions over the membrane. A recent breakthrough study shed light on the biochemical function of RNA degradosome-mediated RNP-bodies (Al-Husini *et al.*, 2020). The

main RNA substrates captured in these degradosome-based RNP bodies are small regulatory RNAs, antisense RNAs and poorly translated mRNAs. rRNAs (ribosomal RNAs) and tRNAs (transfer RNAs), on the other hand, are not enriched in degradosome condensates. Not only do these RNP bodies stimulate RNA decay of target RNAs, they also prevent accumulation of potentially harmful endo-cleaved degradation intermediates by promoting complete mRNA turnover (Al-Husini *et al.*, 2020). In summary, it is apparent the bacterial RNP condensates help organise RNA turnover in the cell. At the basis of these nano-organelles is the RNA degradosome with its intrinsically disordered character and multiple RNA binding sites. That the unstructured regions in the degradosome seem to play key roles in local, transient compartmentalisation may pose a strong additional evolutionary force for these scaffold domains to persist in evolution.

2.3 On-the-fly RNA surveillance opportunities in bacteria

To prevent accumulation of aberrant transcripts that may be both deleterious and a concomitant waste of energy, mRNA quality control mechanisms have evolved in eukaryotes and prokaryotes. Early studies showed that physical coupling of transcription-translation-degradation provides the cell with a means to monitor for non-translated transcripts, terminate their synthesis (Richardson, 1991), and promote their decay (Lost & Dreyfus, 1995). In eukaryotes, cytoplasmic RNA decay happens co-translationally, and the Ski helicase complex is key in this coupled process. In particular, the Ski helicase complex, which assists the cytoplasmic exosome in RNA 3' → 5' decay and quality control pathways, was found to associate with stalled 80S ribosomes. The interaction exposes the 3' end of a mRNA substrate directly off the ribosome. Via the Ski2 helicase channel, the mRNA target is then fed to the exosome for degradation (Schmidt *et al.*, 2016). In eukaryotes, such on-the-fly RNA surveillance mechanisms likely prevent formation of aberrant transcripts.

There is evidence for transcription-coupled translation and translation-coupled RNA degradation in bacteria too. The so-called 'expressome' is an RNA polymerase-30S small ribosomal subunit complex that was recently characterized structurally and shown to exist in *Mycoplasma pneumoniae* by *in situ* tomography studies (Kohler *et al.*, 2017; O'Reilly *et al.*, 2020). It was shown that both transcription and translation were needed for an expressome to be active. In particular, transcription inhibition resulted in stalled expressomes whereas translation inhibition resulted in dissociation of the expressomes (O'Reilly *et al.*, 2020). Recent evidence indicates that the coupling of transcription-translation may not occur in the representative Gram-positive bacterium *Bacillus subtilis*, which may also be the case for other related Firmicutes (Johnson *et al.*, 2020). Instead, other ribonucleoprotein complexes that form co-transcriptionally may play corresponding roles in regulation of translation and decay in these bacterial phylae.

There is precedent for interplay between translation and degradation in bacteria as well, where mRNA lifetime could be influenced mostly by the time during which it can support protein synthesis. In *E. coli*, for example, the RNA degradosome can bind ribosomes during translation, and subsequently degrade the single stranded mRNA template (Tsai *et al.*, 2012). In this scenario, the mRNA is degraded as a consequence of translation inhibition and RNA decay is involved only as a scavenging process (Dreyfus, 2009; Deana & Belasco, 2005). In these studies, the

degradosome associated DEAD-box helicase RhlB is predicted to directly interact with the stalled ribosome, analogous to the Ski-complex in eukaryotic RNA surveillance complexes. Interestingly, RNase J based degradosomes have also been shown to associate with translating ribosomes in *Helicobacter pylori* and *B. subtilis* (Redko *et al.* 2013). In addition, a *B. subtilis* endoribonuclease with a PIN domain was recently found to associate with the ribosome A site and cleave nascent transcripts in a ribosome dependent mRNA decay pathway (Leroy *et al.*, 2017). It should be noted, however, that besides degradation of a pool of target transcripts the ribosome associated RNA degradosomes could play a role in rRNA maturation as well, but this is yet to be elucidated. Like in eukaryotes, the interactions described above could contribute to complex RNA surveillance mechanisms in bacteria, where the RNA degradosome prevents aberrant transcripts from being expressed, and at the same time rescues ribosomal assemblies.

2.4 Rationale and research aims

Many studies, both functional and structural, have been carried out on individual components or small sub-assemblies of the RNA degradosome. However, no structural research has been published so far on larger sub-assemblies or the entire complex. It is clear that studying the degradosome as a whole is very challenging and many years of research are required to further develop reproducible approaches to do so. This chapter aims to contribute to these developments by describing purification strategies, tailored SEC-SAXS ensemble approaches and a tomography/sub-tomogram averaging pipeline to study the RNA degradosome. In particular, methods were developed to study the three research themes described in the introduction: 1. The natively unstructured character of the degradosome; 2. Membrane association of the RNA degradosome and 3. Interactions between the RNA degradosome and ribosomes or ribosomal subunits.

3 Results

3.1 Expression and purification of degradosome assemblies

Reproducible expression, reconstitution and purification of large RNA degradosome assemblies has been one of the main bottlenecks for structural and functional studies. This is in part due the membrane association of the degradosome and in part due to its natively unstructured and exposed scaffold region, rendering it prone to protease cleavage during purification. In this first section, reproducible purification strategies were developed for a truncated version of the RNA degradosome as well as the full RNA degradosome.

3.1.1 *In vivo reconstitution* of the **truncated** RNA degradosome.

The truncated degradosome is a tetrameric 945 kDa complex comprising RNase E (1-850), RhlB and Enolase (Figure 2D). The rationale for using a truncated version of the RNA degradosome stems from the possibility that the trimeric PNPase, containing three RNase E interaction sites, organises degradosome assemblies into a continuous membrane-associated network in the cell by connecting C-terminal scaffold domains (Górna *et al.*, 2012). To prevent this from happening in solution and to ensure discrete assemblies, RNase E was truncated to 850 amino acids, removing the PNPase binding site altogether. Some of the experiments described below were performed in the presence of RNA substrate in the sample. To ensure that the RNA substrate binds but is not degraded, catalytically inactive mutants of RNase E (1-850) were cloned (Materials and Methods section 5.1). In particular, a D303R/D346R mutant was prepared, as these mutations have been shown to render the RNase E catalytic domain inactive (Callaghan *et al.*, 2005). Later on, these mutations were changed to the more stable but equally inactive D303N/D346N substitutions (NN-truncated degradosome). For the truncated degradosome activity assays described in chapter II, wild type truncated degradosome was used (WT-truncated degradosome). Finally, a Manganese-inducible version of RNase E (D346C) was also cloned (Mn-truncated degradosome). The latter was cloned for research done by Dr. Saiful Islam, but not used for experiments described in this chapter.

The same purification strategy was used for all constructs described above. The truncated degradosome components were co-expressed in *E. coli* ENS134-10 cells. *E. coli* ENS134-10 cells have the endogenous full length *rne* gene, encoding RNase E, substituted for *rne1-850* to prevent chimeric RNase E/RNase E (1-850) tetramers from forming during heterologous expression. An N-terminal His-tag was fused to RNase E (1-850), whereas RhlB and Enolase were untagged. Unfortunately, the transformation efficiency after co-transforming multiple plasmids was poor, which was narrowed down to a very low transformation efficiency for the Enolase encoding pET21b plasmid. Therefore, pET21b was transformed into *E. coli* ENS134-10 separately, after which the cells were made chemo competent again via standard procedures (*E. coli* ENS134-10_eno). For all subsequent purifications, *E. coli* ENS134-10_eno cells were transformed with a pRSF-DUET plasmid encoding RNase E (1-850) and RhlB. Cell growth and expression of the truncated degradosome is described in detail in Materials and Methods section 5.2.1.

Before lysing the ENS134-10 cells, a series of additives were included to solubilise the degradosome and prevent proteolysis of the RNase E scaffold domains by endogenous proteases. 1% (v/v) Triton X-100 and 0.1% (w/v) ß-

DDM (final concentrations) were used for degradosome solubilisation. EDTA-free protease inhibitor tablets (Roche, according to the manufacturer's protocol) and 1mM PMSF (final concentration) were added to prevent protease digestion. Lastly, 1mM TCEP (final concentration) and 2 µg/ml of DNase I were added to the cells to prevent oxidation and to cleave unwanted DNA species during cell lysis. For all subsequent purification steps, 0.02% ß-DDM was added to the buffer to keep the degradosome soluble. After lysis by cell rupture and centrifugation a Nickel column purification step was carried out to pull down RNase E (1-850). This purification step was carried out in a high salt background (1M NaCl) to prevent co-purification of degradosome bound RNA species. Next a cation exchange purification step was carried out with a SP column to further remove RNA impurities from the sample. In a final purification step, aggregated complexes were removed from the sample with a Superose 6 10/300 GL size-exclusion column (Figure 2A). All purification steps were carried out at 4°C and within 24 hours, to limit proteolysis as much as possible. Using this highly reproducible purification strategy, the resulting truncated degradosome samples are pure and with very limited proteolysis contaminants (Figure 2A). From 6 x 0.5L 2xYT cultures 5mg of truncated degradosome was purified on average. For all truncated degradosome experiments in this chapter, the NN-truncated degradosome was used, unless stated otherwise.

3.1.2 Pull down of the endogenous **full** RNA degradosome

Several approaches were taken to purify the whole *E. coli* RNA degradosome (see also sections 3.1.3 and 3.1.4). Initially, attempts were made to purify the endogenous degradosome from *E. coli* NCM3416 cells (Carpousis lab, LMGM, University of Toulouse, France), which carry a chromosomal strep-tagged RNase E (N-terminal). A detailed protocol is described in Materials and methods (section 5.2.2). The same additives were added to the NCM3416 cells before lysis as for the truncated degradosome, but the lysate was loaded on a Strep column in a high salt background and eluted with Desthiobiotin. Peak fractions were pooled and loaded onto a Heparin column. The degradosome was eluted with a linear gradient of a high salt buffer. The resulting samples were relatively clean in terms of protein content, with only a few weak contaminant bands on an SDS PAGE gel (Figure 2B), and the purified endogenous RNA degradosome was shown to be active (data not shown). However, the yield was low, with only 0.01 mg of protein from 10 x 1L cell cultures, and the purified sample was contaminated with nucleic acid species. Therefore, an *in vitro* reconstitution protocol for the full RNA degradosome was developed (section 3.1.3).

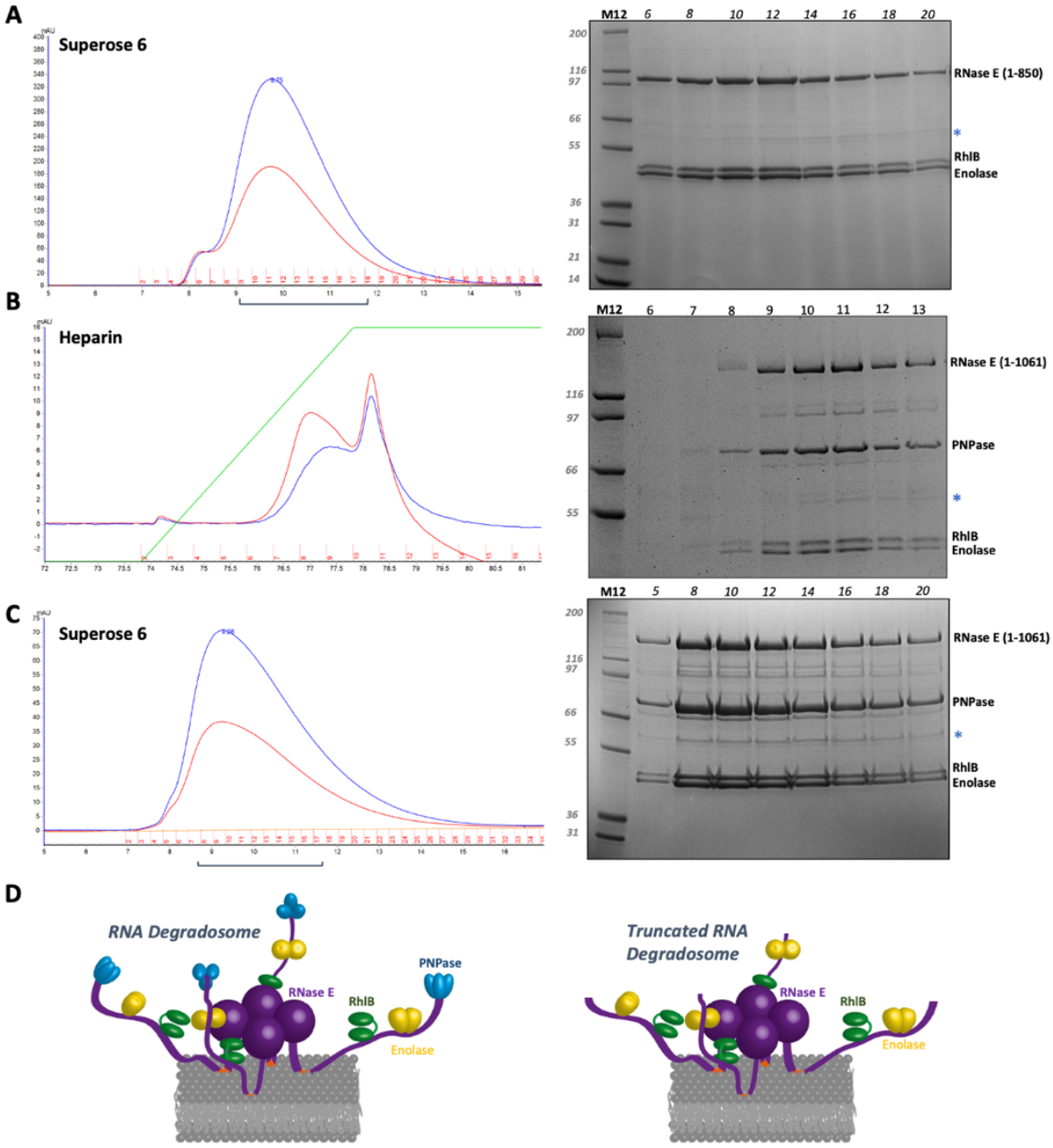


Figure 2: ***Overview of RNA degradosome purification results. (A)*** *Gel filtration elution profile of the in vivo reconstituted truncated degradosome (left) and subsequent SDS-PAGE analysis (right).* ***(B)*** *Heparin elution profile of the endogenous RNA degradosome (left) and subsequent SDS-PAGE analysis (right) after trichloroacetic acid (TCA) precipitation. No further analysis was performed on the endogenous RNA degradosome.* ***(C)*** *Gel filtration elution profile of the in vitro reconstituted RNA degradosome (left) and subsequent SDS-PAGE analysis (right). Fractions used for further experiments are highlighted by an accolade under the elution profiles. Protein standards are annotated on the left of each gel (in kDa). Protein bands are annotated on the right of each gel, corresponding fraction numbers are annotated on top of each gel. M12, Mark 12 protein ladder. * indicates a known proteolysis product, i.e. the RNase E catalytic domain.* ***(D)*** *Schematics of the RNA degradosome (left) and the truncated degradosome (right). Panel D was adopted from Bandyra et al., 2013.*

3.1.3 *In* vitro reconstitution of the **full** RNA degradosome

To achieve a higher yields and lower RNA contamination in the full RNA degradosome sample, an *in vitro* reconstitution protocol was developed. Both wild type RNase E and a D303N/D346N mutant were cloned and purified. Components of the degradosome were expressed separately. Enolase, PNPase and RhlB were expressed separately in *E. coli* BL21(DE3) cells. 6xHis-tagged RNase E (1-1061) (pET15) was expressed in *E. coli* C43 cells. Cells harvested separately from 6 x 0.5 L RNase E culture, 1 x 0.5 L Enolase culture, 1 x 0.5 L PNPase culture and 1 x 0.5 L RhlB culture were mixed together during high pressure cell rupture to reconstitute the RNA degradosome *in vitro*. The next steps are identical to the purification protocol for the truncated degradosome (section 3.1.1). All steps were performed at 4°C and the purification was done within 24h to prevent extensive proteolysis of the RNase E scaffold domain. Nevertheless, some contaminating bands and proteolysis products remained after three purification steps (Figure 2C). From the mixed culture volumes described above, 0.5-1 mg of full-length RNA degradosome was purified.

3.1.4 *In vivo* reconstitution of the **full** RNA degradosome

Attempts were made to reconstitute the full RNA degradosome *in vivo*. Co-transforming four different plasmids (one for each RNA degradosome component) for *in vivo* assembly was unsuccessful. Therefore, only 6xHis-RNase E (1-1061) was expressed in *E. coli* and purified as described in section 3.1.1. Interestingly, endogenous PNPase co-purified at stoichiometric amounts (data not shown). RhlB and Enolase, however, co-purified at sub-stoichiometric amounts (data not shown). Therefore, this *in vivo* assembly approach was abandoned.

3.2 The truncated degradosome is a discrete, homogeneous assembly

After optimising expression and purification protocols for the RNA degradosome, experiments were set up to probe the oligomeric state of the purified degradosome, i.e. tetrameric or monomeric RNase E, and explore its natively unstructured character. The analyses entailed analytical ultracentrifugation (AUC), dynamic light scattering (DLS) and SEC-SAXS (Size Exclusion Chromatography-Small Angle X-ray Scattering). For all these experiments the ribonuclease inactive NN-truncated degradosome, for which sufficiently high yields were routinely achieved, was used.

3.2.1 Analytical ultracentrifugation

Analytical ultracentrifugation (AUC) is a rigorous and accurate biophysical technique that is used to estimate the molecular weight and homogeneity of a protein or protein assembly and to assess its stability (Cole *et al.*, 2008). A sample is typically subjected to high centrifugation forces inside an AUC cell and the sedimentation velocity or sedimentation equilibrium inside the cell is measured by tracking UV absorbance or interference of the sample along the cell. To calculate the sedimentation coefficient, truncated degradosome was freshly prepared at 0.4 mg/ml in 50 mM Tris-HCl pH 7.5, 400 mM NaCl, 100 mM KCl, 5 mM $MgCl_2$, 1 mM TCEP, 0.02% ß-DDM and 10% (v/v) glycerol. The sedimentation velocity of the truncated degradosome was measured over 12 hrs, while spinning at 4°C in vacuum, 50 000 g. Analysis of the AUC data was carried out in SedFit (Zhao *et al.*, 2013) and resulted in an estimated mass of ~833 kDa after solving the Lamm equation (Figure 3A). The buffer conditions and temperature were corrected for by calculating the buffer density and viscosity with SEDNTERP (http://jphilo.mailway.com). According to the AUC data, the sample is homogeneous, with 85% of the particles corresponding to the main peak. Interestingly, a frictional ratio of 1.48 was estimated, indicating that the truncated degradosome is extended. The theoretical mass for the truncated degradosome is 945 kDa, which is within the 10-15% accuracy limits of AUC, indicating that RNase E forms homo-tetrameric degradosome particles in solution.

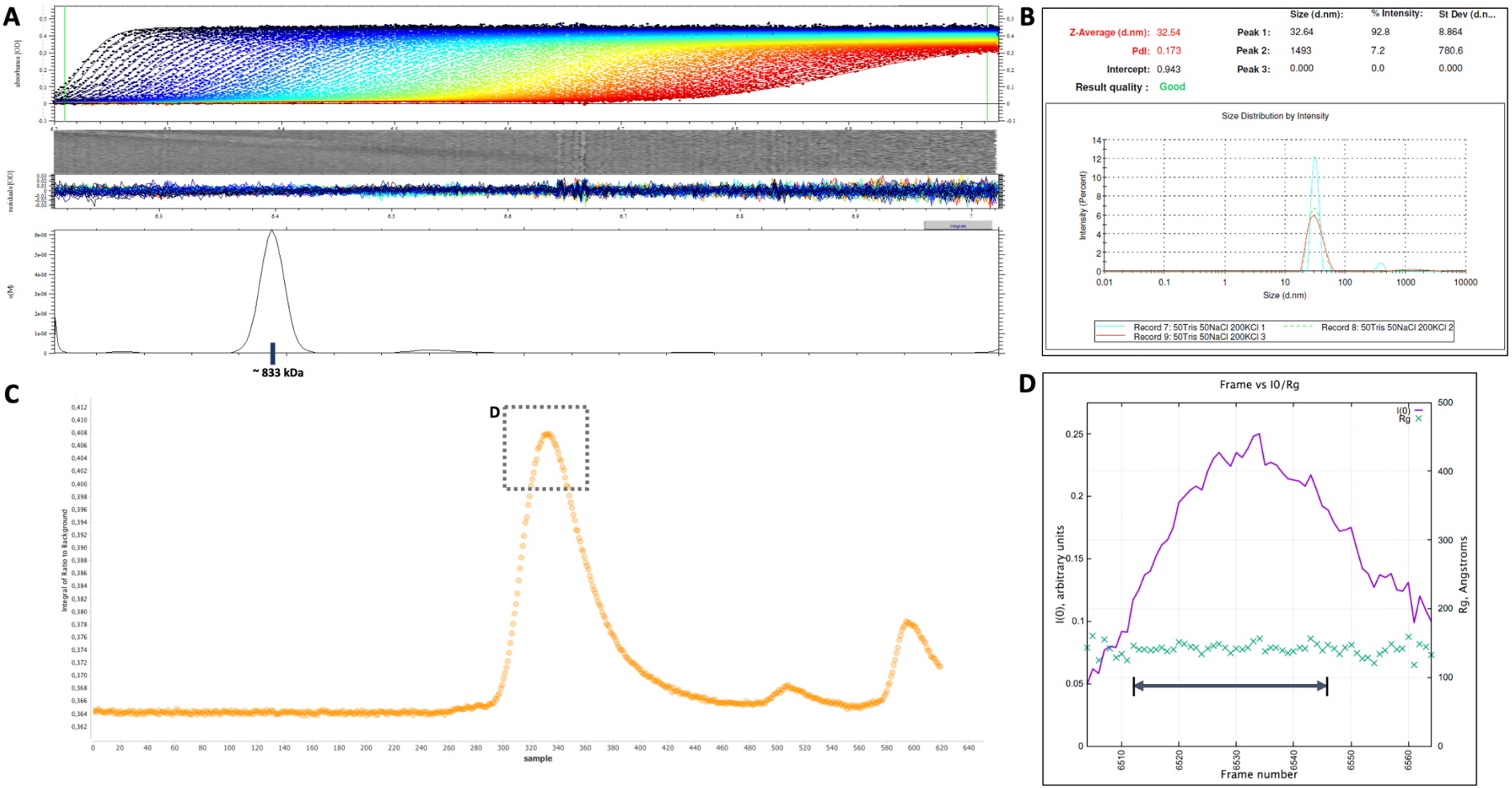


Figure 3: ***Biophysical analyses of the truncated RNA degradosome. (A)*** *Analytical ultracentrifugation experiments reveal a homogenous truncated degradosome with a molar mass estimated at 833 kDa. A frictional ratio of 1.48 was estimated, indicating that the truncated degradosome is elongated. Absorbance profiles during the ultracentrifugation are shown in the top panel, fitting residuals are shown in the middle panel and molar mass distribution is shown in the bottom panel.* ***(B)*** *Dynamic light scattering experiments indicate that the truncated degradosome is homogenous and has a hydrodynamic size of 325 Å. DLS experiments were performed in triplicate.* ***(C)*** *The SEC profile from SEC-SAXS experiments shows that the truncated degradosome elutes as*

a single peak. ***(D)*** *The Radius of gyration, R_g, as monitored over the elution peak with dataSW (Shkumatav & Strelkov, 2015), shows that the truncated degradosome assembly has a high degree of compositional homogeneity. The double arrow marks the frames that were used to generate the final intensity curve.*

3.2.2 Dynamic light scattering

Dynamic light scattering was used as a secondary check for truncated degradosome homogeneity and to estimate particle dimensions by calculating its hydrodynamic size. DLS monitors the Brownian motion of particles in solution with light scattering and fits an auto-correlation function to deduce the translational diffusion coefficient, D, which is lower for larger particles. The diffusion coefficient is then used to estimate the hydrodynamic radius, a measure for the size of the particle together with its solvation layer, from the Stokes-Einstein equation:

$$d_H = \frac{kT}{3\pi\eta D}$$

, with d_H the hydrodynamical radius, k the Boltzmann constant, T the absolute temperature, η the viscosity and D the translational diffusion coefficient.

DLS experiments were carried out in close collaboration with Ms. Giulia Paris. Figure 3B shows a representative outcome of a DLS experiment on the truncated degradosome in an intermediate-salt buffer (50 mM Tris-HCl pH 7.5, 200 mM KCl, 50 mM NaCl, 5 mM $MgCl_2$, 1 mM TCEP and 0.02% ß-DDM). Three measurements were taken per degradosome sample, and size distributions were plotted by intensity. The truncated degradosome sample is monodisperse, with a polydispersity index (PDI) recorded at 0.173, and has a hydrodynamic size estimated at approximately 330 Å in the buffer above (Figure 3B). The latter again indicates that the truncated degradosome is an extended protein complex.

3.2.3 SEC-SAXS

To further explore the structural characteristics of the truncated degradosome, Small Angle X-ray Scattering coupled to Size Exclusion Chromatography (SEC-SAXS) experiments were carried out. In SAXS experiments, a solution of the protein of interest is exposed to monochromatic beam of X-rays and the scattered X-rays are recorded (see thesis Preface for a brief introduction to SAXS). This X-ray scatter profile, or the intensity profile, I(q), can be used to extract important structural information such as oligomeric state, flexibility, compactness, homogeneity and overall shape of a protein or protein assembly (Skou *et al.*, 2014). Coupling SAXS to Size Exclusion Chromatography has allowed for the separation of oligomeric states/aggregation in the sample prior to exposure to X-rays and has been shown to significantly improve reproducibility (Mathew *et al.*, 2004).

Here the purified truncated degradosome was concentrated to 2.2 µM, 4.4 µM, 7.2 µM and 14.4 µM for SEC-SAXS measurements at the I22 beamline at Diamond light source (Harwell campus, Didcot, UK). The samples were loaded on a Superose 6 Increase 3.2/300 GL equilibrated in 50 mM Tris-HCl pH 7.5, 400 mM NaCl, 100 mM KCl, 5 mM $MgCl_2$, 1 mM TCEP and 0.01% ß-DDM. For all samples loaded the degradosome eluted as a single, reasonably symmetric peak (Figure 3C). As the sample eluted from the column the radius of Gyration, R_g (Å), of the truncated degradosome was monitored over the peak (Figure 3D). This confirmed that the purified truncated degradosome assembly was indeed monodisperse.

The most information-rich output from a SAXS experiment is the one-dimensional intensity profile (I(q)), which plots the intensity of the scattered X-rays as a function of resolution. For SEC-SAXS, these intensity curves are typically obtained by radially averaging the circular scattering patterns for each frame in the SEC elution profile (Figure 3C) and subtracting the intensity profile of the running buffer from them. Next a suitable range of frames under the elution peak is defined and the intensity profiles for these frames are averaged together (Figure 3D). The resulting intensity curve forms the basis for all subsequent analyses.

3.2.3.1 Structural characteristics of the truncated RNA degradosome

The averaged intensity curves for the different truncated degradosome samples are presented in Figure 4A. The higher the concentration of truncated degradosome in the sample, the better the signal to noise ratio, especially at regions corresponding to higher spatial frequency (high q range). All four curves have the same overall shape and align well (Figure 4A). At low spatial frequency (low q range), however, higher concentrations correlate non-linearly with I_0 values (I_0 is the intensity at q=0), and the intensity curves have a relatively poor overlap (Figure 4B). These observations are indicative of interparticle interactions in the sample at higher concentrations. Given the perfect overlap at intermediate and high q ranges, the intensity curves were merged with ALMERGE (Figure 4C) (Franke *et al.*, 2012). In this merged intensity curve the low spatial frequency range was taken from that of the sample with the lowest concentration (2.2 μM), free of inter-particle interactions, and the intermediate/high q ranges was taken from the most concentrated sample (14.4 μM).

Overall the intensity profile for the truncated degradosome is rather featureless (Figure 4C). This indicates that the curve is likely to represent an average of different intensity profiles, each corresponding to a different conformational state of the truncated degradosome in solution. In other words, this observation is a first strong indication that the truncated degradosome is conformationally heterogeneous, and that the experimental intensity curve represents a pool of widely different conformational states (Figure 6A). Guinier analysis shows that there are no signs of aggregation in the sample, and estimates the Radius of gyration (R_g), which is a measure for the overall spread of mass in the particle, at 137 Å (Figure 4D). Even though the protein assembly is close to 1 MDa in molecular weight, such a large R_g indicates that the truncated degradosome is a highly extended protein assembly. The indirect Fourier transform of the intensity curve I(q) generates the pair-distribution function (P_r), which is the SAXS equivalent of the Patterson function in X-ray crystallography. The P_r is a distance distribution function of all atom pairs in the protein assembly and is used to e.g. estimate the maximum inter-atomic distance (D_{max}). For the truncated degradosome the D_{max} was estimated at 427 Å, which again indicates that the degradosome is highly extended (Figure 4D). Moreover, the P_r function is bimodal, which is typical for multidomain assemblies. Finally, the Kratky plot (normalised for radius of gyration) indicates that the truncated degradosome has a high degree of flexibility (Figure 4E). In particular, the shape of the Kratky plot, with a large 'secondary bump', is a signature for multidomain assemblies with flexible linkers.

All of the observations above confirm, for the first time, the prediction that the truncated degradosome and likely the full RNA degradosome are extended assemblies that are conformationally heterogeneous in solution.

Therefore, no low-resolution *ab initio* reconstructions of the truncated degradosome shape were generated, as these would represent an average shape of many conformations and would contain limited information. Instead, an integrative structural biology approach was used in section 3.2.4 to perform ensemble optimisation modelling. This strategy aims to describe and quantify the degree of conformational heterogeneity in the truncated degradosome.

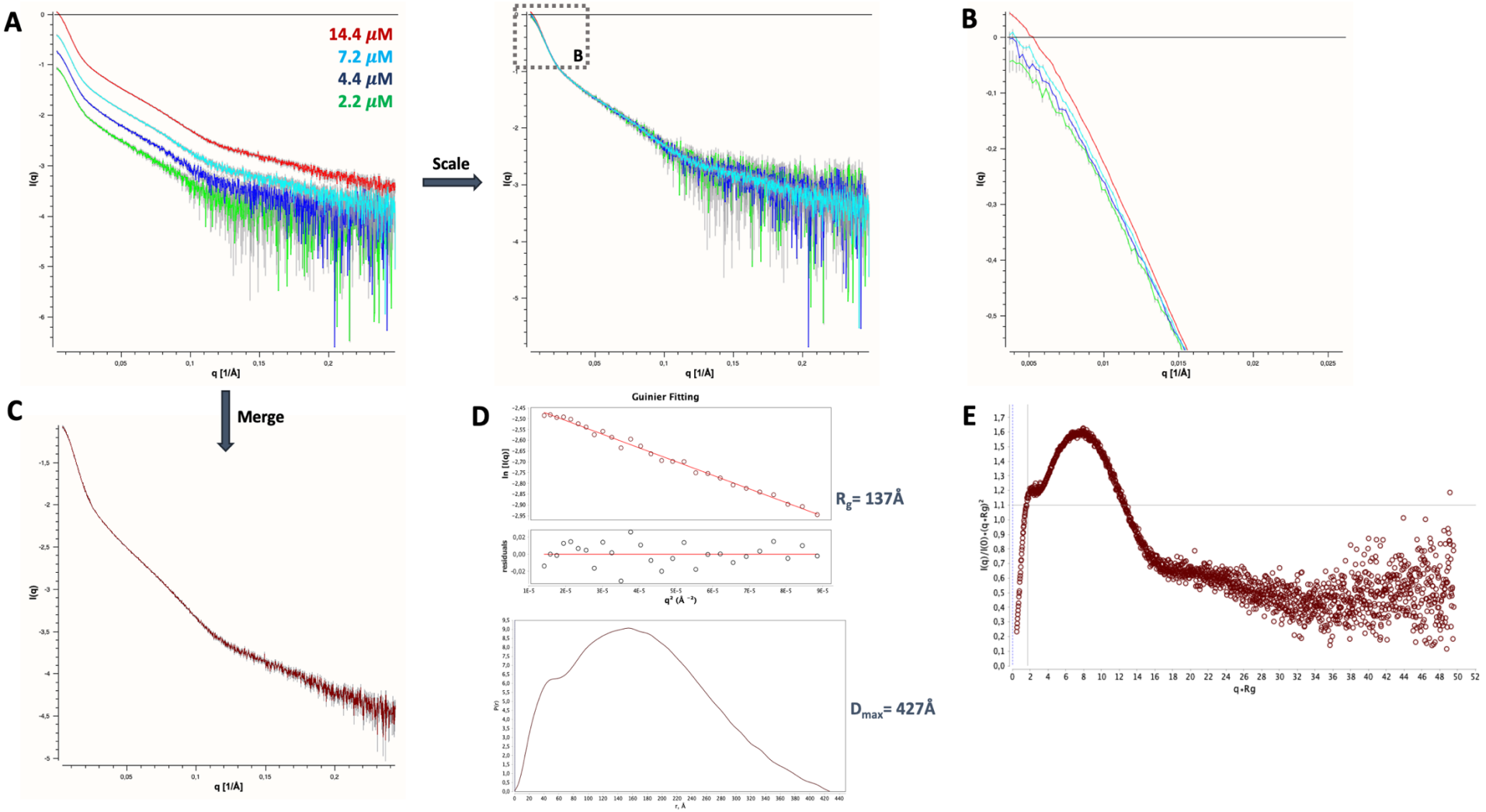


Figure 4: ***SEC-SAXS analysis of the truncated degradosome.*** ***(A)*** *Raw and scaled intensity profiles for the four truncated degradosome samples tested. Concentrations are annotated and colour coded.* ***(B)*** *At low spatial frequency (low q-range) the intensity profiles do not align due to mild interparticle interactions.* ***(C)*** *A merged intensity profile was generated using the low q-range of the truncated degradosome sample with the lowest concentration, and the intermediate/high q-range of the intensity profile for the 14.4 µM sample.* ***(D)*** *Guinier analysis results in a radius of*

gyration, R_g, of 137 Å, and the pairwise distribution function extends to a D_{max} of 427 Å. ***(E)*** *The Kratky plot (normalized for R_g) indicates that the truncated degradosome is a multidomain protein assembly with flexible linkers.*

3.2.3.2 *The truncated degradosome compacts upon RNA binding*

Next, a sample was prepared for SEC-SAXS experiments of the NN-truncated degradosome supplemented with 9S rRNA at equimolar ratios. 9S rRNA is a highly structured precursor of the 5S ribosomal RNA and is processed by RNase E in *E. coli* (Ghora & Apirion, 1978). The overall shape of the resulting intensity curve is similar to the curves for the truncated degradosome alone (Figure 5A). However, Guinier analysis reveals that the truncated degradosome is more compact when 9S rRNA is bound, with a R_g of 123 Å (137Å for the truncated degradosome alone), despite the additional molecular weight. The pairwise distribution function also reveals that the truncated degradosome assembly is comparatively less extended in the presence of 9S, with a D_{max} of 374 Å (427 Å for the truncated degradosome alone) (Figure 5B). Moreover, the normalised Kratky plot (normalised for R_g) shows that the NN-truncated degradosome (red) becomes more rigid with 9S rRNA binding (blue) (Figure 5C). This is apparent from the less pronounced 'secondary bump' in the Kratky curve. These observations indicate that the truncated degradosome becomes more compact when bound to substrate RNAs, such as the 9S rRNA.

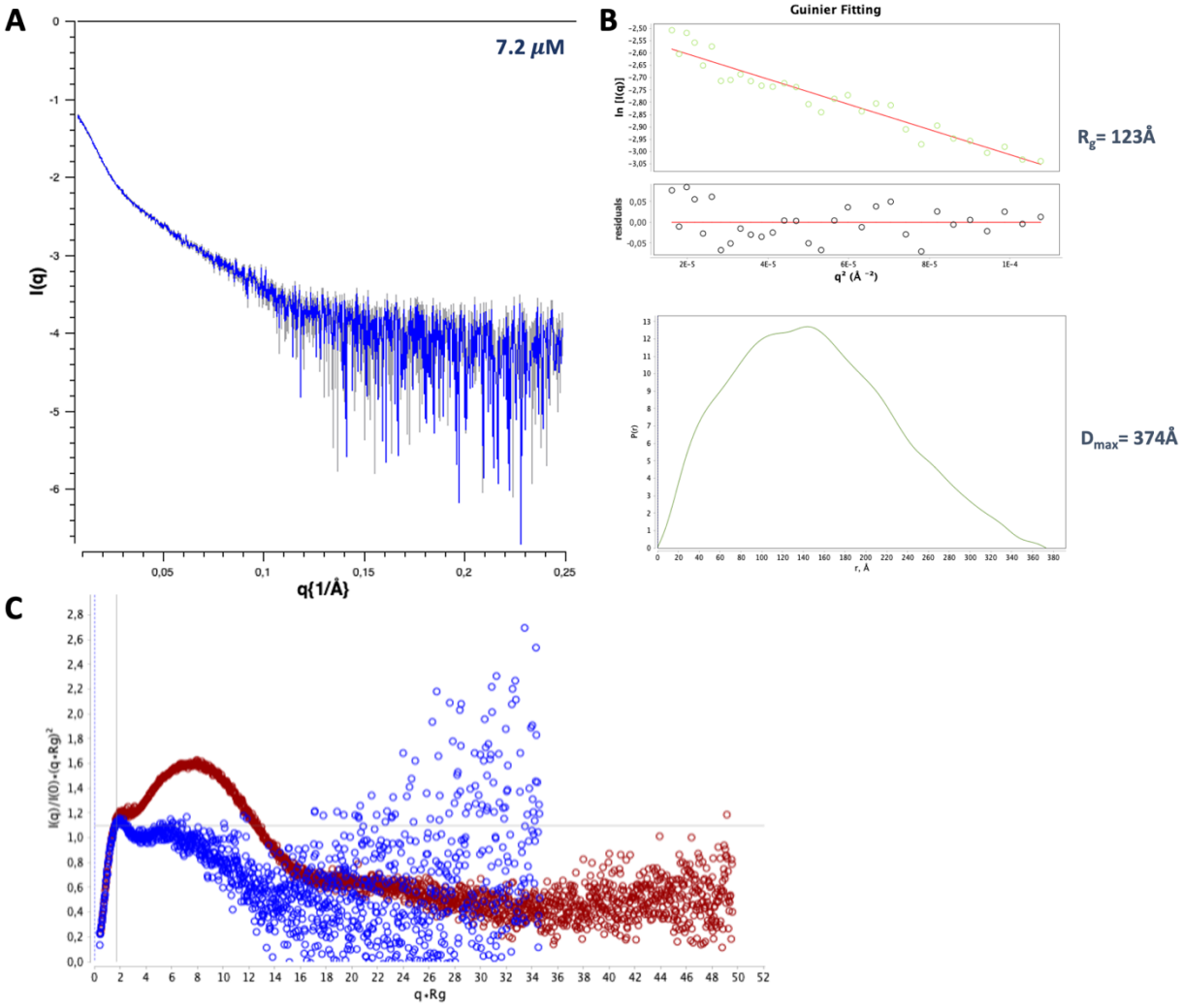


Figure 5: ***SEC-SAXS analysis of the truncated degradosome- 9S rRNA complex. (A)*** *Radially averaged intensity profile for the truncated degradosome bound to 9S rRNA.* ***(B)*** *Guinier analysis and the pairwise distribution function suggest that the 9S-truncated degradosome assembly is less extended, with a R_g of 123 Å and a D_{max} of 374 Å.* ***(C)*** *The normalized Kratky plot further*

confirms that the complex is more compact in the presence of RNA. The blue circles correspond to the 9S-truncated degradosome complex and the red circles correspond to the truncated degradosome by itself.

3.2.4 Conformational behaviour of the truncated degradosome, an integrative structural biology approach

SEC-SAXS is a powerful biophysical tool that is often used to describe the structural behaviour of a protein or protein complex. For the truncated degradosome, SEC-SAXS experiments show the protein assembly is highly extended with respect to the molecular mass and flexible, but that it somewhat compacts upon RNA substrate binding. The flexibility limits the information that can be extracted from raw SEC-SAXS intensity curves for the degradosome using classical approaches (Figure 4). In this section, published structural data on components of the RNA degradosome were used in an integrative structural biology approach to quantify/visualize its inherent conformational heterogeneity. For this, a tailored pipeline was developed based on previously published Ensemble Optimisation Modelling (EOM) approaches (Bernado *et al.*, 2007; Tria *et al.*, 2015) (Figure 6). Intuitively, EOM tries to explore the conformational landscape of proteins in solution based on *a priori* structural information and experimental intensity curves. The approach uses a genetic algorithm to extract ensembles of models that, as a mixture, explain the SAXS data from a large pool of candidate structures, generated based on prior knowledge, thus attempting to define physical boundaries around the conformational possibilities (Figure 6). Since the number of ensemble solutions for a typical SAXS curve is astronomically large, regardless of the degree of prior knowledge, this technique cannot resolve the structure of the truncated degradosome, but can find ensembles of structures for which the distribution of physical properties (e.g. R_g and D_{max}) resembles that of the true ensemble in solution.

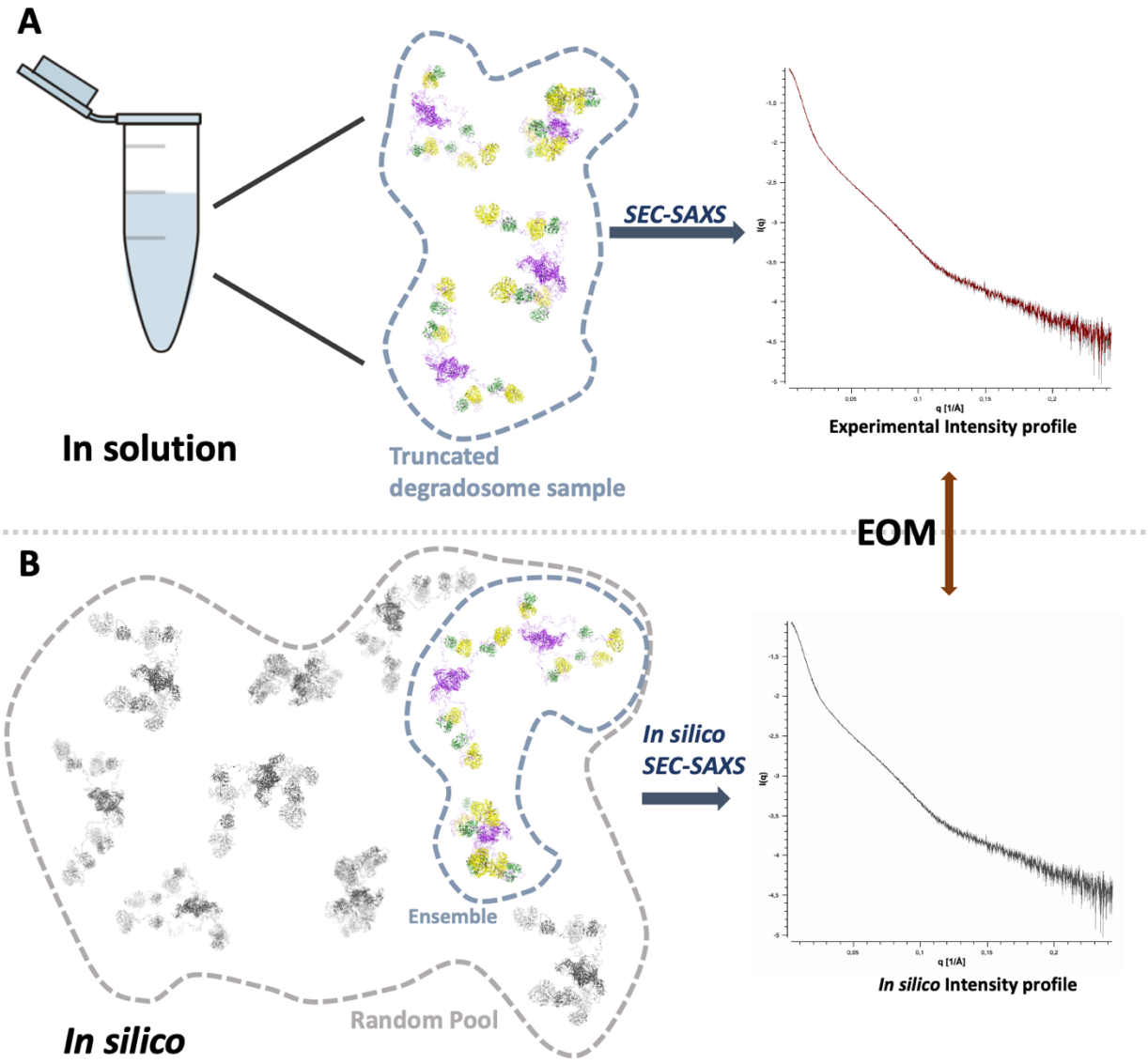


Figure 6: ***Schematic overview of the Ensemble Optimization Modelling approach (EOM). (A)*** *The SAXS curve for the truncated degradosome represents the scattering data of an ensemble of conformationally heterogeneous assemblies.* ***(B)****. EOM starts from a large in silico pool of random truncated degradosome structures and searches for ensembles that explain the experimental intensity curve. To do this, in silico intensity curves are generated for each selected ensemble with CRYSOL (Svergun et al., 1995) and compared to the experimental intensity curve for the truncated degradosome.*

The bottleneck for applying EOM on the truncated degradosome is the generation of a random pool of candidate structures. The EOM software was written for single chain proteins with flexible domains, whereas the truncated degradosome is a multiprotein assembly. Therefore, a tailored EOM pipeline was developed with the help of Dr. Aliaksandr Shkumatau (Efremov group, VUB, Brussels, Belgium). The classical EOM strategy consists of two steps, each with their own algorithms. RANCH generates a random pool of structures, and takes prior structural information as an input, and GAJOE uses a genetic algorithm to perform ensemble modelling on the random pool, against experimental SAXS data (Bernado *et al.*, 2007) (Figure 7). For all components of the RNA degradosome there is extensive structural data available that can be used in the form of restraints for RANCH. First, a crystal structure of the RNase E catalytic domain in its tetrameric apo-form was solved by Koslover *et al.* (2008) (pdb-ID

2VMK). Second, Enolase has been co-crystalized with a fragment of the RNase E scaffold domain by Nurmohamed *et al.* (2009), revealing a structural basis for the interaction interface between these two (pdb-ID 3H8A). Finally, two studies have elucidated the interaction between RNase E and RhlB (Chandran *et al.*, 2007; Bruce *et al.*, 2018). No structure has been solved for RhlB from *E. coli* so far, but an accurate homology model based on VASA helicase from *Drosophila* has been generated in prior studies (Sengoku *et al.*, 2006; Bruce *et al.*, 2018). Here, CABS-Dock (Blaszczyk *et al.*, 2016) was used to model the interaction between RhlB and a fragment of RNase E (as found by Chandran *et al.*, 2007) with constraints based on the HDX-MS experiments (Hydrogen Deuterium Exchange – Mass Spectrometry) performed by Bruce *et al.* (2018) (Figure 7A). Taken together, the only parts of the truncated degradosome that are unaccounted for, and therefore will be sampled via EOM, are the RNase E scaffold domains (Figure 7B and C).

Since RANCH can only generate pools of single chain proteins, RhlB and Enolase were made 'invisible' to RANCH by converting them to 'hetero atom' models (see Materials and Methods section 5.6 for script). Using the structure for a protomer of the RNase E catalytic domain, the crystal structure of Enolase bound to a fragment of the RNase E scaffold domain and the RhlB homology model bound to a fragment of the RNase E scaffold domain, a random pool of truncated degradosome protomers was generated with RANCH, where RNase E was modelled as a C-α trace model (Figure 7B). To generate these protomeric truncated degradosome models, RANCH modelled the RNase E scaffold domain as a 'random coil' for each candidate structure in the protomer pool, thereby sampling conformational space (Figure 7B). The same procedure was repeated for each protomer of the tetrameric RNase E catalytic domain, generating four 'monomer' pools of 20 000 protomers each. It is important to note that for each model the RNase E catalytic domain was fixed in space, which will be referred to in the next paragraph. Finally, FULCHER (Shkumatau *et al.*, in press.) was used to convert all truncated degradosome protomers in the four 'monomer' pools from invisible 'hetero atom' C-α models back to visible 'all atom' models, resulting in a pool of monomeric truncated degradosome protomers (Figure 7B).

From the four pools of protomeric truncated degradosome structures (Figure 7B), random combinations were picked to generate a 'tetramer' pool of 10 575 160 candidate truncated degradosome structures with CombinerIM (script written by Dr. Aliaksandr Shkumatau) (Figure 7C). By fixing the positions of the RNase E catalytic domain in space, as described above, the tetrameric RNase E core was always reconstituted when combining four truncated degradosome protomers, one from each 'monomer' pool. Besides generating random tetrameric truncated degradosome structures, CombinerIM also uses phenix.clashscore to calculate the model clash score and subsequently filter out models with a clash score above the user defined threshold (Liebschner *et al.*, 2019). For the truncated degradosome, it was found that a clash score threshold of 60 was sufficiently stringent to remove sterically forbidden candidate structures. Of the initial pool of 10 575 160 truncated degradosome structures, 2 920 079 were retained after clash score filtering. Lastly, CombinerIM uses CRYSOL to calculate an *in silico* intensity profile for each member of the filtered truncated degradosome pool (Figure 7D) (Svergun *et al.*, 1995). This pool of *in silico* intensity profiles is then used as input for GAJOE, which uses a genetic algorithm to select

ensembles of structures/intensity profiles from the random pool that explain the experimental SAXS curve for the truncated degradosome (Figure 6B and Figure 7E).

GAJOE was run for 100 rounds on the filtered tetramer pool, composing an ensemble of conformations in each round. Each selected ensemble comprised of 4-7 degradosome models, with goodness of fit (χ^2) scores between 3.6 - 3.8 against the experimental intensity curve. Next, distribution functions were plotted for the R_g and D_{max} metrics of the selected ensembles and compared to the corresponding distribution functions for the random pool (Figure 8A and B). From these distribution functions it is clear that the ensemble models are significantly more extended than the overall pool models. In other words, the RNase E scaffold domains of the truncated degradosome extend outwards from the catalytic core in solution. These observations are in agreement with quantitative metrics based on the size distributions, such as the 'flexibility' (R_{flex}) and the 'variance' (R_σ). R_{flex} is an entropy-based estimation of the flexibility of a pool of models, where R_{flex}= 0% indicates complete rigidity and R_{flex}= 100% corresponds to maximal flexibility. R_σ, on the other hand, is the ratio between the ensemble standard deviation and the random pool standard deviation, i.e. how wide a range of model sizes are populated by the ensembles compared to the random pool (Tria *et al.*, 2015). For the truncated degradosome R_{flex}(ensemble) = 79.01% and the flexibility within the random pool is nearly identical (R_{flex}(pool) = 78.95%). Furthermore, R_σ was calculated at 1.01, meaning that the model sizes in the selected ensembles are as variable as those in the random pool (yet shifted to significantly higher values). Together with the size distributions, these values indicate that the ensemble modelling did not find a set of rigid truncated degradosome models that explain the SAXS data, i.e. the truncated degradosome adopts many different conformations in solution, all of which are highly extended (Figure 8A, B and D). The best ensemble ((χ^2 = 3.6, Figure 8C) consists of four extended truncated degradosome models and is depicted in Figure 8D.

Notably, the D_{max} distribution of the ensembles has three maxima, which suggests that clusters of conformers could exist in solution (Figure 8B, grey arrows). The rule of thumb for reliable annotation of such sub-populations is that their maxima are separated by at least two times the standard deviation of the random pool (Tria *et al.*, 2015). This is the case for the second and third maxima in the D_{max} distribution of the ensembles which suggests that truncated degradosome conformers can be grouped in distinct clusters of extended and somewhat less extended assemblies in solution.

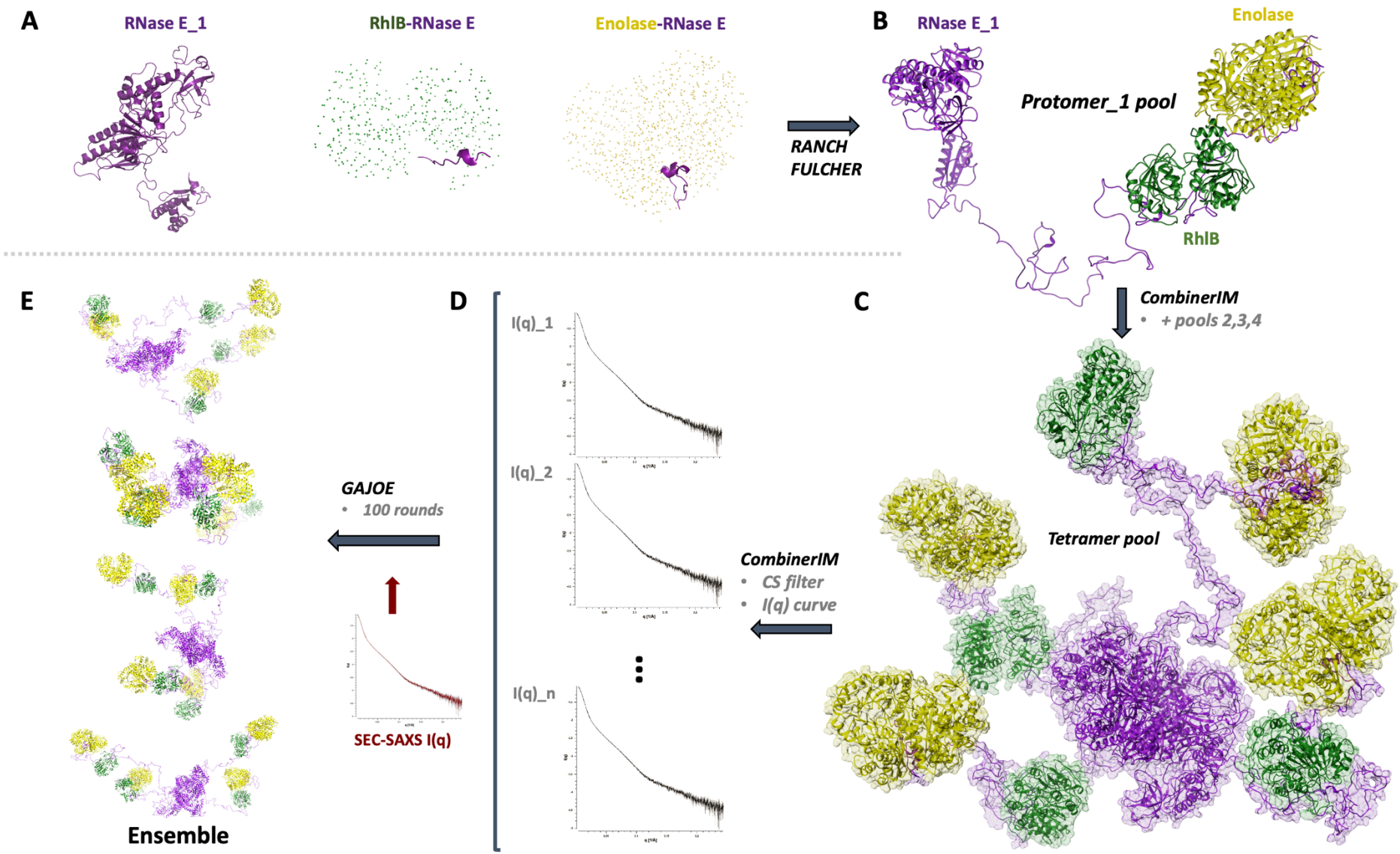


Figure 7: ***Schematic overview of the tailored Ensemble Optimization Modelling (EOM) pipeline used for the truncated degradosome. (A)*** *Prior knowledge about truncated degradosome components and their interactions with the RNase E scaffold domain was combined in the form of a crystal structure (RNase E, purple), a co-crystal structure (Enolase, yellow) and a homology*

model (RhlB, Green). RhlB and Enolase were made 'invisible' to RANCH and are presented as dots. ***(B)*** *RANCH was used to generate random truncated degradosome protomers based on (A), modelling the RNase E scaffold domain as an intrinsically disordered peptide. FULCHER was used to convert Enolase and RhlB back to all-atom models. For each truncated degradosome protomer a random 'monomer' pool of 20 000 structures was generated.* ***(C)*** *CombinerIM was used to combine protomers of each pool to generate random structures of tetrameric truncated degradosome assemblies. As such, a pool 10 575 160 random truncated degradosome structures was generated.* ***(D)*** *Next, CombinerIM was used to filter the random truncated degradosome pool based on a clash score threshold (CS filter) and calculate in silico intensity profiles for each assembly with CRYSOL (Svergun et al., 1995).* ***(E)*** *This final pool of 2 920 079 intensity profiles was scanned with GAJOE to find ensembles that, when averaged together, fit the experimental intensity profile for the truncated degradosome.*

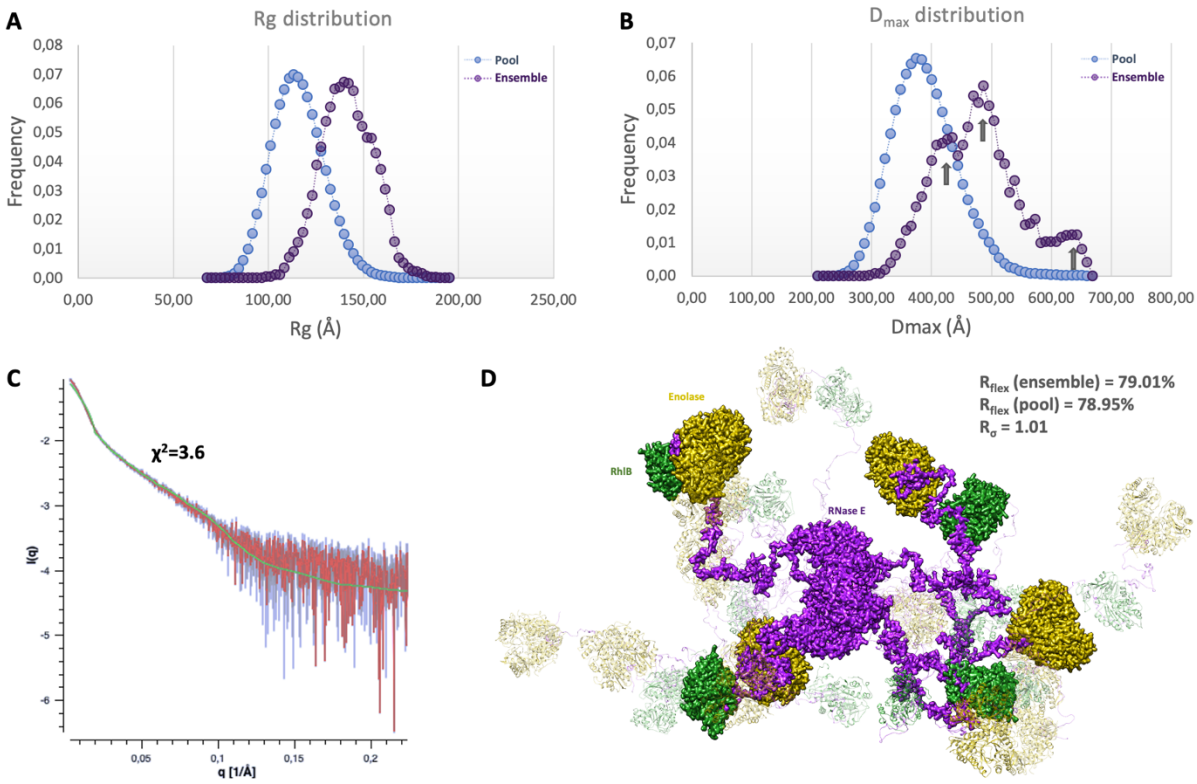


Figure 8: ***Overview of the EOM results. (A)*** *Distribution of the radii of gyration for the random pool (blue) and the ensembles (purple) of the truncated degradosome. Highly extended truncated degradosome conformers were selected by GAJOE.* ***(B)*** *Distribution of D_{max} values for the random pool (blue) and the ensembles (purple) of the truncated degradosome. The different maxima in the ensemble distribution could correspond to distinct clusters of truncated degradosome conformers (grey arrows).* ***(C)*** *Intensity curve of the best ensemble (green) fitted to the experimental SAXS curve (red) (χ^2 = 3.6).* ***(D)*** Overlay of the four models in the best ensemble (color coded), illustrating the extendedness of the truncated degradosome.

3.3 Electron microscopy studies of the truncated degradosome

The SEC-SAXS experiments above revealed that the truncated degradosome is a very flexible protein assembly. In particular, the individual enzyme components of the RNA degradosome are likely to behave as loose beads on long, natively unstructured strings, i.e. the RNase E scaffold domain. Interestingly, the truncated degradosome compacts somewhat upon binding substrate RNAs such as 9S rRNA. In the remainder of this chapter, conditions will be explored, most of them biologically relevant, that might further change the flexible behaviour of the truncated degradosome and organize its components in a more defined way. These approaches include stabilisation by cross-linking (3.3.1), interactions with translation machinery (3.3.3) and interactions with a lipid membrane (3.3.4). Most of the experiments described below are electron microscopy based. Recent advances in image processing for electron microscopy allow to explore and model substantial conformational heterogeneity in protein assemblies (Nakane *et al.*, 2018; Punjani & Fleet, 2020; Zhong *et al.*, 2020). These approaches all start from a stack of particle images with a relatively accurate set of consensus alignments, which results in an interpretable 3D reconstruction. If a low-resolution consensus map of the truncated degradosome could be reconstructed, despite the inherent conformational heterogeneity, these new algorithms could be used to further explore and model its flexibility. Therefore, different electron microscopy and sample preparation approaches were tested to find out if 2D images of the truncated degradosome would align.

3.3.1 GraFix stabilises the truncated degradosome

A first approach was to test whether the truncated degradosome forms discrete and interpretable particles on an EM grid, either for negative stain EM (continuous carbon grids) or cryo-EM (holey carbon grids) (Figure 10A and B, left panel). The truncated degradosome was expressed and purified as described in section 3.1.1 and diluted to 5-100 µg/ml for negative stain studies or 0.3-0.5 mg/ml for cryo-EM studies. Negative stain grids were prepared by pipetting the sample onto glow discharged continuous carbon grids, blotting away excess sample manually and staining with 1% (w/v) uranyl acetate (Materials and Methods section 5.11.1). The grids were screened on a FEI Tecnai F20 (200kV) (Department of Material Sciences, University of Cambridge) (Figure 10A, left panel). No distinct particles were visible on the grid, indicating that the truncated degradosome aggregated and/or dissociated during grid preparation. In parallel, cryo-EM grids were prepared with a Vitrobot MarkIV (FEI) (Materials and Methods section 5.11.2) and screened on a FEI Talos Arctica (200kV) (BioCem, Department of Biochemistry, University of Cambridge). Here too, truncated degradosome complexes were aggregated and appeared like diffuse heterogeneous assemblies (Figure 10B, left panel).

To stabilise the truncated degradosome particles before grid preparation, the GraFix protocol was optimised and employed (Kastner et al*., 2007*; Stark, 2010). GraFix is an ultracentrifugation technique coupled to mild chemical cross linking (Figure 9) and was briefly mentioned in chapter II. The protocol is described in detail in the Materials and Methods, section 5.8. In short, 200 pmol of truncated degradosome was applied on top of a glycerol/glutaraldehyde gradient in an ultra-centrifugation tube. The tubes were spun in an ultracentrifuge for 18 hours to remove aggregates/dissociated complexes and crosslink the sample (Figure 9A). The gradient was then

fractionated from bottom to top (Figure 9B). Fractions containing protein were detected via a standard dot blot and used for grid preparation.

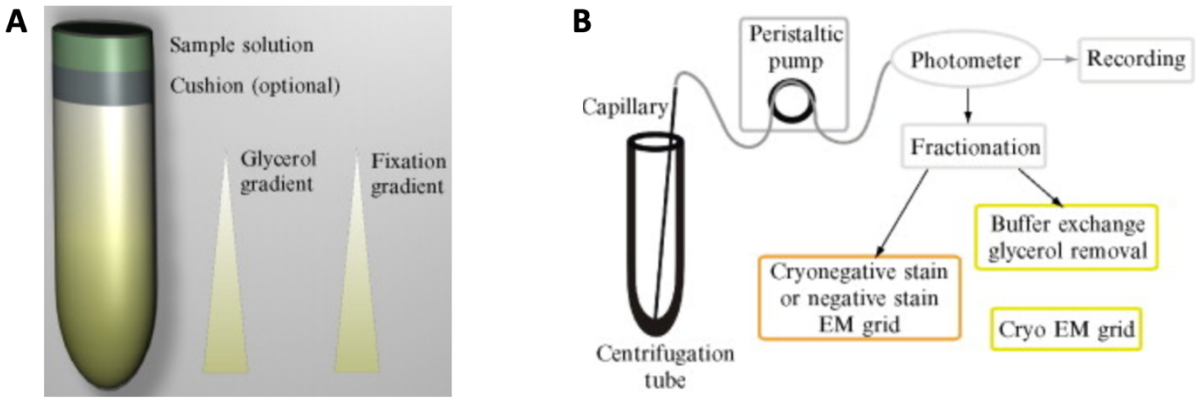


Figure 9: ***Overview of the GraFix strategy. (A)*** *A continuous glycerol and crosslinker gradient were set up in an ultracentrifugation tube and the sample was applied to the top of the tube.* ***(B)*** *After ultracentrifugation, the gradient was fractionated with a peristaltic pump (bottom to top) and protein-containing fractions were analysed via negative stain EM or via cryo-EM after buffer exchange to remove the glycerol. Images are adopted from Stark, 2010.*

For negative stain studies the same grid-preparation protocol was used as described in the previous paragraph. The screening images showed that the GraFix samples behaved significantly better on the grid, with distinct and reasonably homogeneous particles of 20-30 nm in size (Figure 10A, middle panel). A small dataset of 30 images was collected and processed in Relion-2.1 (Scheres *et al.*, 2012). 1500 particles were picked manually and classified in 2D (Figure 10A, right panel). From these 2D class averages it was apparent that, after GraFix, the crosslinked truncated degradosome is a compact, spherical particle. GraFix was also used to prepare samples for cryo-EM analysis. A dialysis step was included after fractionation of the density gradient, to remove glycerol from the sample, which was subsequently concentrated down to 20 µl (from 600 µl). Cryo-EM grids were prepared the same way as described above. Here too the screening images revealed a clean sample and reasonably homogeneous particles of 25-30 nm, albeit at low abundance (Figure 10B, middle panel). To test whether the particles aligned, a dataset was collected on a 300 kV Titan Krios (FEI) (Nanoscience Centre, University of Cambridge). 1400 movies were processed in Relion 2.1 (Scheres *et al.*, 2012). Unfortunately, multiple rounds of 2D classification resulted in poorly resolved class averages, some of which had a somewhat triangular shape (Figure 10B, right panel). Further 3D classifications did not result in a reproducible 'consensus' reconstruction, even at low resolution (data not shown). It is likely that the crosslinking reaction during GraFix generated many different intraparticle crosslinks, which compacts the truncated degradosome but does not necessarily ameliorate conformational heterogeneity. Finally, RNA substrates, like 9S rRNA, were added to the sample prior to grid freezing, in an attempt to make the truncated degradosome more compact and rigid. These attempts were not successful and resulted in aggregated clumps of protein in the gird holes (data not shown).

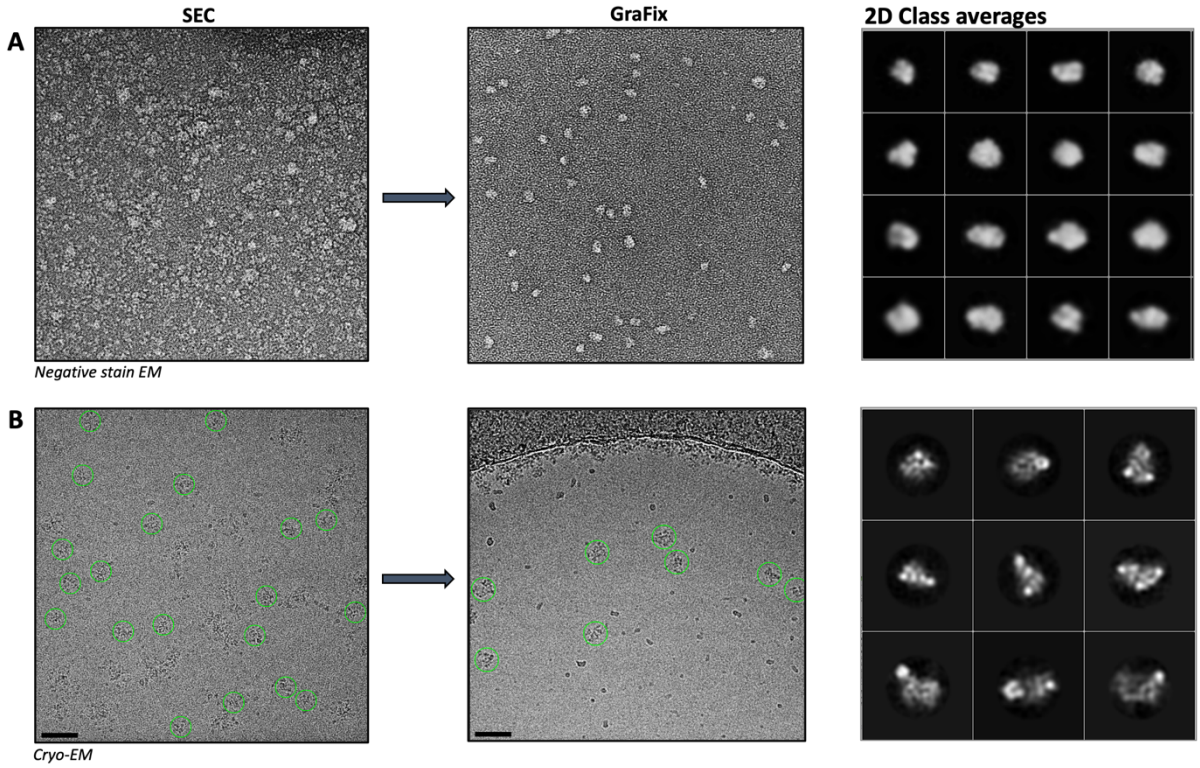


Figure 10: ***Overview of negative stain EM and cryo-EM studies. (A)*** *Conventional size exclusion chromatography steps (SEC) did not yield uniform, well resolved particles on the uranyl acetate stained grids (left). GraFix was used to remove dissociated truncated degradosome components and aggregates, yielding well resolved particles (middle). 2D classification in Relion 2.1 resulted in compact, spherical, but featureless class averages (right).* ***(B)*** *Conventional size exclusion chromatography steps (SEC) did not yield uniform and well-resolved particles on cryo-EM grids (left). After GraFix, however, distinct and compact particles were visible in the grid holes (middle). 2D classification rounds in Relion 2.1 resulted in poorly resolved, somewhat triangular-shaped class averages.*

3.3.2 The catalytic core of RNase E is not the culprit for poor degradosome behaviour in SPA

Although the AUC, DLS and SEC-SAXS results show that the truncated degradosome is an intact tetrameric assembly (section 3.2), it is possible that the complex dissociates during grid freezing. Often proteins and protein assemblies associate with the air-water interface during grid preparation, which can lead to complete or partial denaturation (Noble *et al.*, 2018; D'Imprima *et al.*, 2019). To test whether the N-terminal domain of RNase E, i.e. the catalytic core of the RNA degradosome, stays intact during grid freezing, RNase E (1-529) was purified (Materials and Methods section 5.2.4) and studied by Cryo-EM. RNase E (1-529) encompasses only the catalytic core of RNase E and was crystallised for the first time by Callaghan *et al.* (2005). Follow-up studies have shown that the tetrameric catalytic unit can undergo significant conformational reorganisations when binding and degrading RNA substrates (Koslover *et al.*, 2008, Bandyra *et al.*, 2018). If the catalytic RNase E core is indeed dynamic in solution, then this could also add to the conformational heterogeneity of the RNA degradosome. In summary, Cryo-EM studies of RNase E (1-529) will reveal (1) whether the RNase E catalytic domain stays intact during grid freezing and (2) whether it is conformationally heterogeneous. Most importantly, if cryo-EM studies of RNase E (1-529) result in an interpretable consensus map, then it must be the RNase E scaffold domains and their binding partners that impede SPA cryo-EM studies of larger degradosome assemblies, either due to their inherent conformational heterogeneity or due to partial denaturation at the air-water interface.

Classical grid preparation approaches on different grid types (e.g. holey carbon and holey gold grids) were unsuccessful, resulting in aggregated clumps of protein in the vitreous ice (data not shown). In chapter II, PNPase was found to align with the air-water interface (AWI) and denature partially at its KH-S1 portal (chapter II, section 3.3.1). The catalytic core of RNase E also has an S1 domain (and a KH-domain) which could align with the AWI in a similar way, with denaturation and/or aggregation of the tetramer as a result. Therefore, it was tested whether CHAPSO, a zwitter-ionic detergent that prevents PNPase from adhering to the AWI, would have the same effect for RNase E. Indeed, when adding 8 mM CHAPSO to the RNase E samples prior to grid freezing, homogeneous particles were observed in the grid holes (Figure 11A). A small dataset was collected on a FEI Titan Krios (BioCem, Department of Biochemistry, University of Cambridge) and processed in Relion 3.0. 2D classification of 180 000 picked particles resulted in well resolved class averages, with apparent C2 symmetry (Figure 11B). After several rounds of 3D classification, a clean subset of roughly 30 000 particles was used for 3D refinement (C1 symmetry),

resulting in a map at 7.8Å resolution (GS-FSC, data not shown). The overall shape of the RNase E cryo-EM map corresponds to the 'open state' crystal structure as solved by Koslover *et al.*, 2008 (PDBid 2VMK) (Figure 11C).

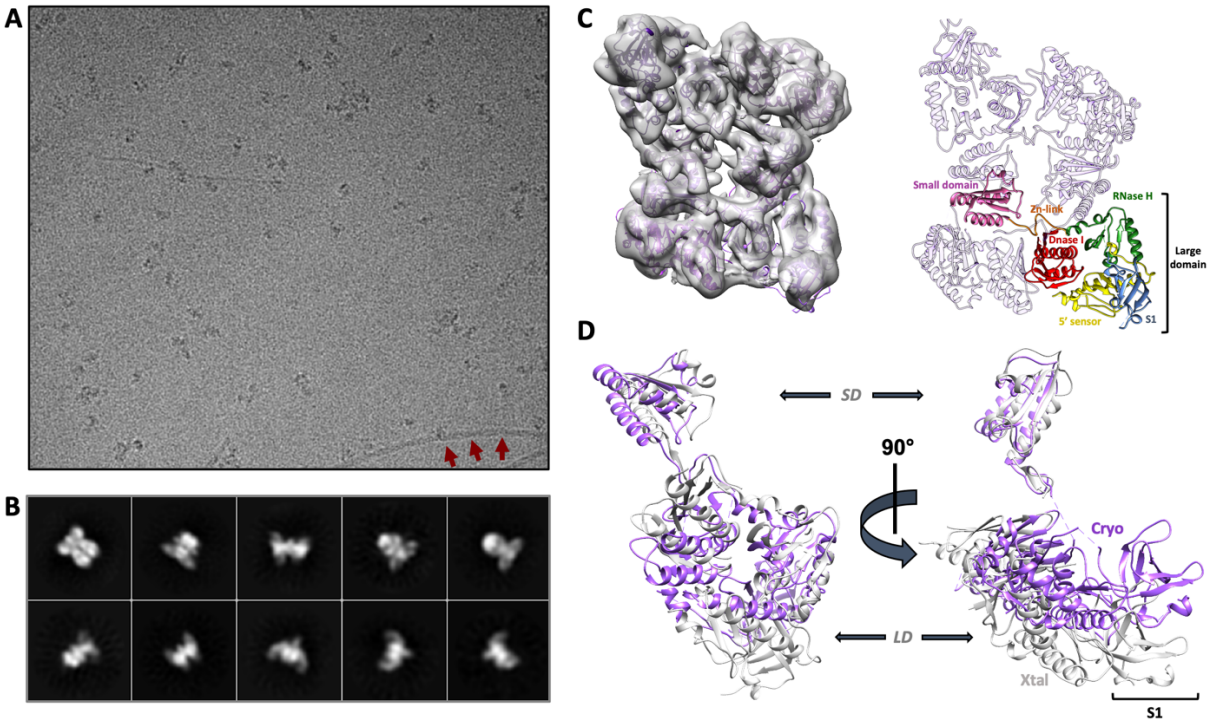


Figure 11: ***Overview of the RNase E (1-529) cryo-EM results. (A)*** *A representative cryo-EM image at 3 µm defocus. Red arrows point to a CHAPSO fibre.* ***(B)*** *2D class averages reveal homogeneous particles with apparent cyclic symmetry (C2).* ***(C)*** *3D refinement of 38K particles results in a 7.8 Å reconstruction of the tetrameric RNase E core (left). A model with colour-coded domains is presented on the right* ***(D)*** *Significant conformational reorganisations are necessary for the RNase E (1-529) crystal structure (pdbID 2VMK) to fit the cryo-EM map. In particular, there is a discrepancy between the RNase E large domains, with a large reorganisation of the S1 domain. Cryo, cryo-EM structure; Xtal, X-ray crystallography structure; SD, small domain; LD, large domain.*

A series of salient observations were made for the cryo-EM results of the RNase E catalytic core. Firstly, the RNase E (1-529) crystal structure required significant conformational reorganisation to fit the cryo-EM map (Figure 11D), the S1 domain in particular. Most likely packing of the RNase E tetramer in the crystal lattice favoured a given conformational state. These packing constraints are not present in vitreous ice, allowing the RNase E tetramer to adopt a more 'relaxed' conformation (Figure 11D). Secondly, the overall resolution of the cryo-EM map is limited, even though the particles are embedded in very thin, vitreous ice, and the image thon rings (in Fourier space) go out to 3Å for most images. Moreover, local resolution estimates reveal an asymmetric pattern in the local resolution distribution, with the top of the RNase E tetramer resolved significantly worse than the bottom (Figure 12A). Lastly, enforcing C2 symmetry during alignments reduces the global resolution, from 7.8Å to 9.5Å (data not shown), even though the crystallographic RNase E (1-529) model is C2 symmetric. The observations described above indicate that there is a significant degree of conformational heterogeneity in the RNase E catalytic core, which impedes high-resolution reconstructions and breaks the overall symmetry of the assembly.

To study potential conformational heterogeneity in the catalytic core of RNase E, a clean set of particles was exported to cryoSPARC to carry out 3D variability analyses (3DVA) (Punjani *et al.*, 2017, Punjani *et al.*, 2020). To account for domain movements that break the symmetry the particles were subjected to C2 symmetry expansion, which copies every particle once and rotates/shifts it according to the symmetry operators and consensus alignments. Since the global resolution does not improve beyond 7.5Å, only conformational variability up to 8Å was considered, and the algorithm was run for 20 iterations, solving six orthogonal modes of variability.

Strikingly, several interpretable modes of conformational heterogeneity were found for the RNase E (1-529) particles (Figure 12B). A first mode of variability describes a synchronous 'rocking' movement of two RNase E large domains (LD) (Figure 12B, Component 1). A second mode of variability describes a 'rolling' movement of all four large domains (Figure 12B, Component 2). Interestingly this rolling mode suggests that an S1 domain can interact with the small domain of a diagonal protomer in the RNase E tetramer (Figure 12B, Component 2, marked with *). The third mode of variability describes a 'breathing' motion of the RNase E tetramer, whereby the assembly expands and contracts synchronously between stable conformations (Figure 12B, Component 3). From these three prevalent modes of variability, it is clear that the catalytic core of the RNA degradosome is itself conformationally heterogenous, albeit to a limited extent, in addition to the flexible scaffold regions. Unfortunately, while adding CHAPSO (8mM) to samples of the larger truncated degradosome resulted in distinguishable particles, these still appeared too heterogeneous for reasonable alignments and reconstructions (data not shown). Finally, it should be noted that the ensemble optimization modelling performed in section 3.2.4 did not account for any conformational heterogeneity of the tetrameric catalytic core of RNase E. However, as the conformational heterogeneity of the catalytic core is moderate compared to that of the RNase E scaffold domains, it is unlikely that it would influence the ensemble modelling outcomes significantly.

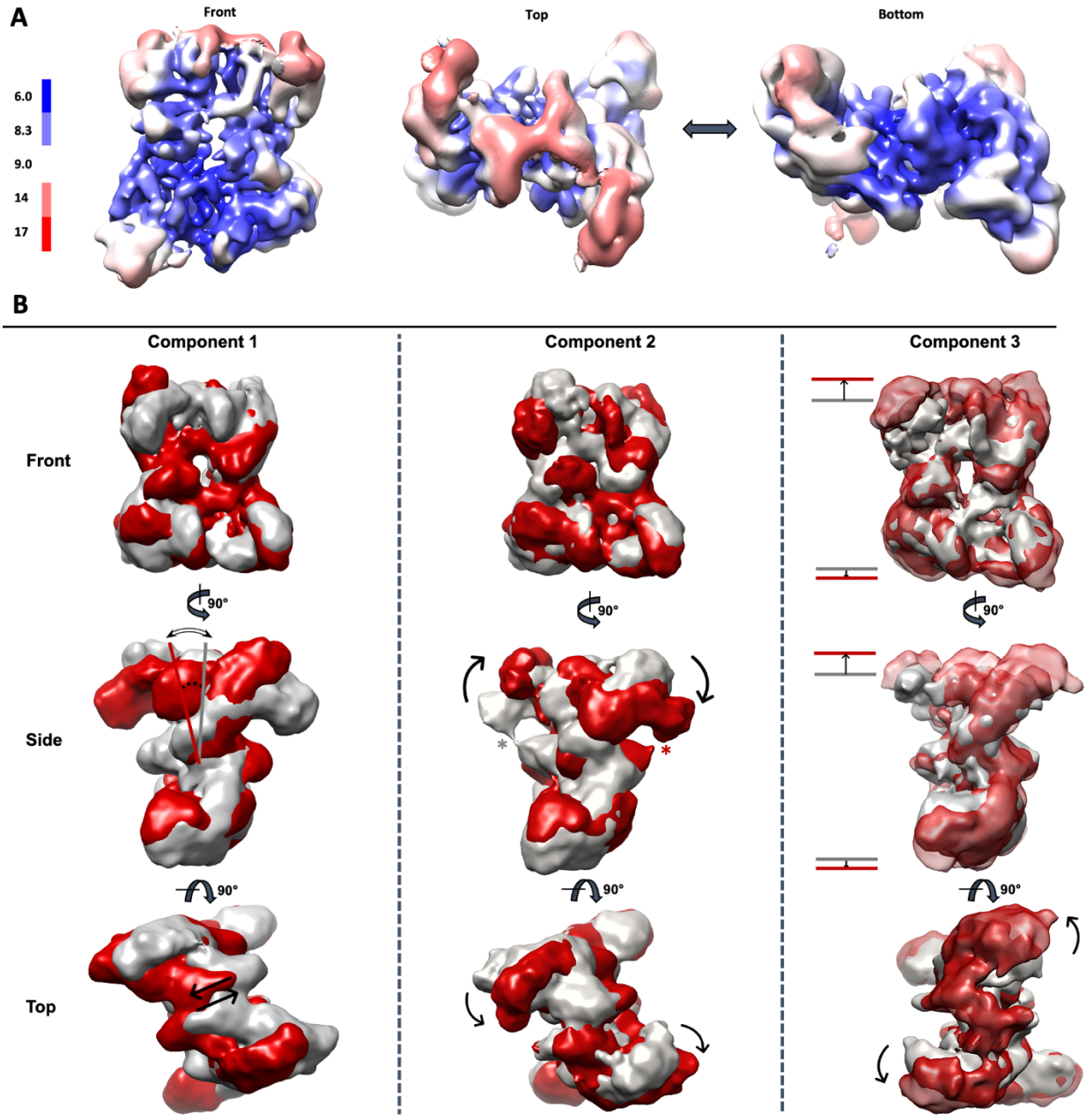


Figure 12: ***Local resolution estimation and 3D variability analysis of the RNase E catalytic core. (A)*** *Local resolution estimation of the cryo-EM map presented in Figure 11C. The top side of the cryo-EM reconstruction is resolved significantly worse than the bottom side.* ***(B)*** *3DVA analysis reveals three main modes of molecular motion within the RNase E tetramer. Variability components 1, 2 and 3 correspond to a 'rocking', 'rolling' and 'breathing' motion, respectively, between low-energy states of the RNase E core. * Highlights a potential interaction between the S1 and small domains of adjacent RNase E protomers (Component 2, side view).*

3.3.3 Interactions between the RNA degradosome and the 30S small ribosomal subunit

The results so far have revealed that direct structural studies of large degradosome assemblies are very challenging, mostly due to its inherent flexible nature. Previous studies have provided evidence that bacterial RNA degradosomes directly interact with translating ribosomes, taking on putative translation surveillance roles (Redko *et al.* 2013; Tsai *et al.*, 2012, Leroy *et al.*, 2017). In *E. coli*, these interactions between the RNA degradosome and ribosomes have been shown to be mediated by RhlB and the two RNA binding sites on the RNase E scaffold domain (RBD and AR2, Figure 1B) (*Tsai et al.*, 2012). It is possible that these interactions help rigidify the degradosome scaffold domains. To test this hypothesis, two RNA degradosome sub-assemblies were used for experiments. One is the truncated degradosome, which is the main assembly studied in this chapter, and the second is the so called 'recognition core' (Bruce *et al.*, 2018). The latter comprises of a 250-residue long fragment of the RNase E scaffold domain (RNase E 603-850), including the RBD and AR2 RNA binding sites, the RhlB binding site and the Enolase binding site, but not the membrane attachment helix (Figure 1B). The truncated degradosome was purified as above, whereas the recognition core, i.e. RNase E (603-850), RlhB and Enolase, were co-purified according to the standard protocol (Bruce *et al.*, 2018). For both these degradosome sub-assemblies, their interaction with the 30S small ribosomal subunit (purified by Prof. Ben Luisi) was investigated using electrophoretic mobility shift assays (EMSA) (section 3.3.3.1). Next, cryo-EM studies were carried on these putative surveillance assemblies (section 3.3.3.2). The EMSA experiments were largely carried out by two master students, Ms. Giulia Paris and Ms. Marta Kubànska, both of whom I supervised.

3.3.3.1 Putative RNA surveillance complexes between the 30S small ribosomal subunit and the RNA degradosome

To test whether the recognition core and the truncated degradosome interact with the 30S small ribosomal subunit, EMSAs were set up in low salt buffer (Materials and Methods section 5.10). Both the recognition core and the truncated degradosome appear to form super-complexes with the ribosomal subunit, as indicated by clear gel shifts (Figure 13A, *). Next, the recognition core and the 30S small ribosomal subunit were co-purified on a Superose 6 10/300 GL gel filtration column for cryo-EM studies (Figure 13B and C).

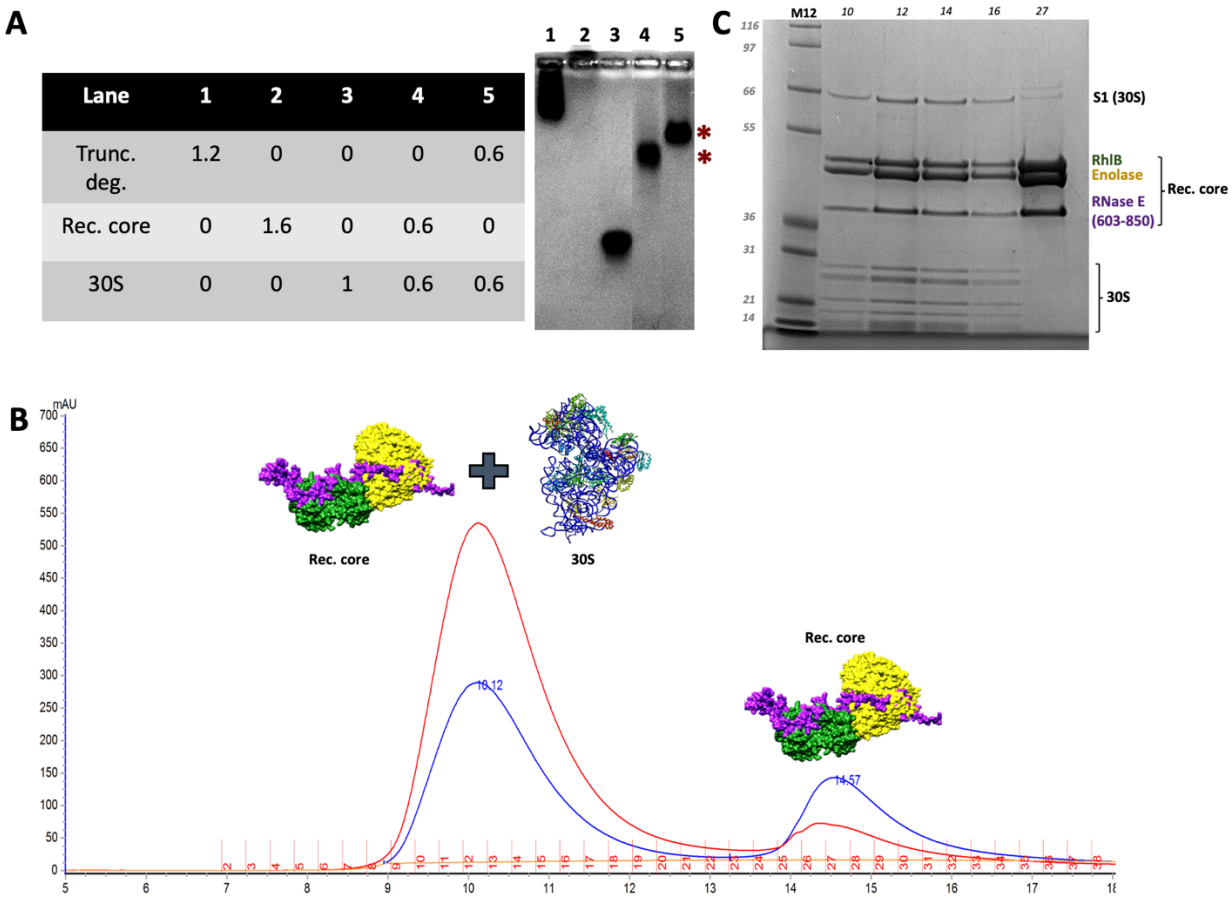


Figure 13: ***EMSA and co-purification of putative surveillance super-complexes. (A)*** *Electrophoretic mobility shift assays (EMSA) indicate that 30S-recognition core and 30S-truncated degradosome super-complexes can form in solution (*). Values in the table are in µM.* ***(B)*** *Co-purification of the 30S-recognition core super-complex via size exclusion chromatography results in two main peaks, one of which has a strong 260 nm absorbance, as expected for the 30S component in the complex. RNase E (603-850) is in purple, RhlB in green and Enolase in yellow.* ***(C)*** *SDS-PAGE analysis of peak fractions reveals that the recognition core co-purifies with the 30S ribosome (lanes 10-16). Fraction 27, corresponding to the second, smaller peak in the elution profile contains excess recognition core. 30S, small ribosomal subunit; Rec. core, Recognition core.*

3.3.3.2 Cryo-EM studies of the 30S-recognition core super-complex

The recognition core and 30S were co-purified as described in the previous section (Figure 13B and C), and concentrated to 0.5 µM prior to grid preparation. Grids were prepared with a Vitrobot Mark IV and screened on a Talos Arctica (FEI, BioCem, Department of Biochemistry, University of Cambridge). 30S particles were readily visible in the cryo-EM images (Figure 14) and a small dataset was collected for the 30S-recognition core super-complex (Talos Arctica) and for 30S alone (negative control, Titan Krios, FEI, BioCem, Department of Biochemistry, University of Cambridge). All image processing and 3D reconstructions were carried out in cryoSPARC (Punjani *et al.*, 2017). Interestingly, diffuse densities were visible near the 30S head in the 2D class averages (Figure 14B, red arrows). Subsequent 3D reconstructions revealed weak additional density near the head/exit channel, which was not observed in the negative control (30S alone). It is likely that this extra density corresponds to components of the recognition core (Figure 14B, purple density). Its location near the mRNA exit site in a corresponding 70S ribosome is conceivable, since this is where the full degradosome would cleave the mRNA during RNA surveillance,

e.g. when a ribosome is stalled during translation. However, only after careful 3D classifications and extensive blurring a map with somewhat reasonable density for the recognition core was obtained. Furthermore, attempts at docking models for Enolase and RhlB were unsuccessful. Crosslinking the assembly with glutaraldehyde was not successful (data not shown). Finally, CHAPSO was added to prevent the 30S subunit from adhering to the air-water interface during grid freezing. Preferred orientation is a known issue for 30S, albeit moderate, and the S1 domains present in the molecule could be responsible for this. Despite the improved orientation distribution of the 30S particles and a much better resolved S1 domain, CHAPSO did not help resolve the recognition core further (data not shown). It is likely that this part of the RNA degradosome still bears significant conformational heterogeneity when bound to the 30S small ribosomal subunit.

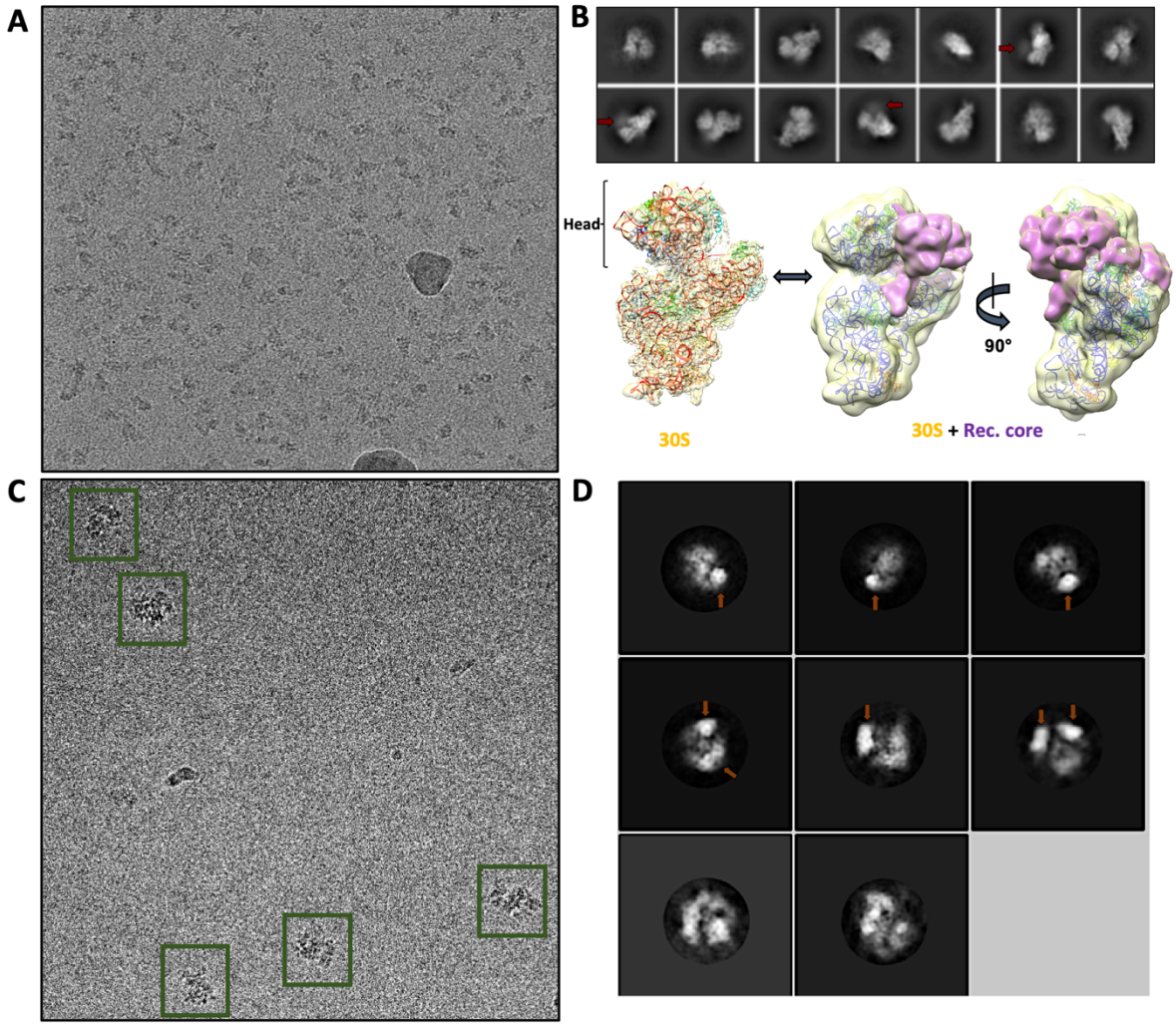


Figure 14: ***Overview of the cryo-EM studies on 30S super-complexes. (A)*** *Representative cryo-EM image of the 30S-recognition core super-complex (2 µm defocus).* ***(B)*** *2D class averages (top) reveal diffuse additional densities near the 30S head region (red arrows). Subsequent 3D reconstructions of the super-complex and comparison to a cryo-EM map of 30S alone reveal weak extra densities spanning across the mRNA exit site (purple).* ***(C)*** *Representative cryo-EM image of the 30S-truncated degradosome super-complex (3 µm defocus). Particles are highlighted with green boxes.* ***(D)*** *2D class averages of the particles reveal high density domains in larger, conformationally heterogeneous assemblies (brown arrows). These domains could correspond to multiple 30S subunits bound to a single truncated degradosome. 30S, small ribosomal subunit; Rec. core, Recognition core.*

Lastly, the 30S subunit and the 945 kDa truncated degradosome were mixed at equimolar ratios, to a final concentration of 0.5 μM, and a small cryo-EM dataset was collected (Talos Arctica, BioCem, Department of Biochemistry, University of Cambridge). Albeit very sparse, large particles were visible in the cryo-EM images (Figure 14C, green boxes). Here too the conformational heterogeneity of the truncated degradosome impeded proper particle alignments with Relion 2.1, and a series of low-resolution 2D class averages was the best result obtained (Figure 14D). Interestingly, the dimensions of these 2D class averages are larger than those of the 30S subunit or even the whole 70S ribosome. Moreover, the 2D class averages appear to have defined domains with increased contrast, which could correspond to the electron dense 30S subunits in the super complex (Figure 14D, brown arrows). If the latter is the case, then multiple 30S subunits could bind a single RNA degradosome assembly, potentially one for each of its four scaffold domains. Unfortunately, this sample of 30S-truncated degradosome super-complexes has been a 'one-hit wonder' so far and attempts to reproduce these images for longer data collections have been unsuccessful.

3.3.4 Structural studies of the membrane bound truncated degradosome

All experiments up to now have been carried out on solubilised RNA degradosome samples in a detergent-rich buffer (0.02% w/v ß-DDM). In this section, protocols were developed to reconstitute the truncated degradosome and full RNA degradosome on lipid membranes, in the form of lipid vesicles, and to study them in a more native environment. The rationale was that since the RNase E membrane anchors are located in the C-terminal scaffold domain, membrane association could lock these in a more rigid conformation (Figure 1B). In addition, the liposomes could help keep the truncated degradosome away from the air-water interface.

3.3.4.1 Reconstitution of the truncated degradosome on lipid vesicles

Lipid vesicles, or liposomes, were prepared by extrusion of a lipid suspension, supplemented with 100 mM sucrose, through a porous membrane (Avanti) (Material and Methods section 5.9). PE (phosphoethanolamine) and PG (phosphatidylglycerol) are the two main phospholipids that constitute the *E. coli* membrane (70% and 20% respectively) (Sohlenkamp *et al.*, 2016). To mimic the *E. coli* plasma membrane, 70% (w/w) DOPE (18:1) was mixed with 30% (w/w) DOPG (18:1). While DOPE is neutral at physiological pH, DOPG is negatively charged. It is important to note that the RNase E membrane attachment helix is amphipathic, with one side of the helix consisting of non-polar amino acids, which embed into the lipid membrane. The other half is positively charged at neutral pH and has been proposed to play a role in facilitating initial interactions with the negatively charged DOPE/DOPG-based *E. coli* membrane (Strahl *et al.*, 2015). To reconstitute the truncated degradosome on the liposomes, 50 µg of truncated degradosome was added to 100 µg of fresh liposomes (200 µl total volume) and dialysed overnight (4 °C) against a detergent-free buffer supplemented with 100 mM glucose instead of sucrose (Materials and methods section 5.9). As such, the truncated degradosome should anchor to the lipid membranes upon depletion of ß-DDM in the sample. After dialysis, the liposomes should have 100 mM sucrose on the interior, making them slightly heavier than the surrounding buffer, which contains 100 mM glucose instead.

To test whether reconstitution of the truncated degradosome on liposomes was successful, ultracentrifugation assays were set up in a glycerol gradient (Figure 15). The protocol is similar to that described for GraFix, but no crosslinker was added to the gradient (Material and Methods section 5.8, Figure 9). 100 µl of a truncated degradosome sample was loaded on top of a 10-40 % continuous glycerol gradient and spun at 84 000 g for 12 hours, 4° C. Since the lipid vesicles are slightly heavier than the surrounding buffer, the centrifugation force should pull them down the glycerol gradient, together with any membrane bound protein assemblies. A peristaltic pump was then used to fractionate the glycerol gradient. Every second fraction was subjected to TCA precipitation and analysed via SDS-PAGE (Figure 15). Three negative control experiments were set up in parallel (Figure 15A-C). Figure 15A shows the ultracentrifugation profile of the truncated degradosome (100 µl at 0.5 mg/ml) in the presence of 0.02% ß-DDM but no liposomes, which indicates that the solubilised truncated degradosome migrates almost halfway down the glycerol gradient. Figure 15B shows the ultracentrifugation profile of the truncated degradosome after dialysing out ß-DDM, in the absence of liposomes. Here too, the truncated degradosome migrates almost halfway down the glycerol gradient, and no large aggregates are observed at the bottom of the tube. Next, Figure 15C shows the ultracentrifugation profile of the truncated degradosome in the presence of

0.02% ß-DDM and liposomes. As expected, the truncated degradosome does not associate with the lipid vesicles when solubilised by ß-DDM and migrates halfway down the glycerol gradient. However, when ß-DDM is depleted from the truncated degradosome sample by dialysis, and liposomes are supplemented, the truncated degradosome adheres to the lipid membrane of the lipid vesicles, as evidenced by the ultracentrifugation profile in Figure 15D. In the latter, the liposome-degradosome bodies migrate almost to the bottom of the glycerol gradient during the ultracentrifugation step.

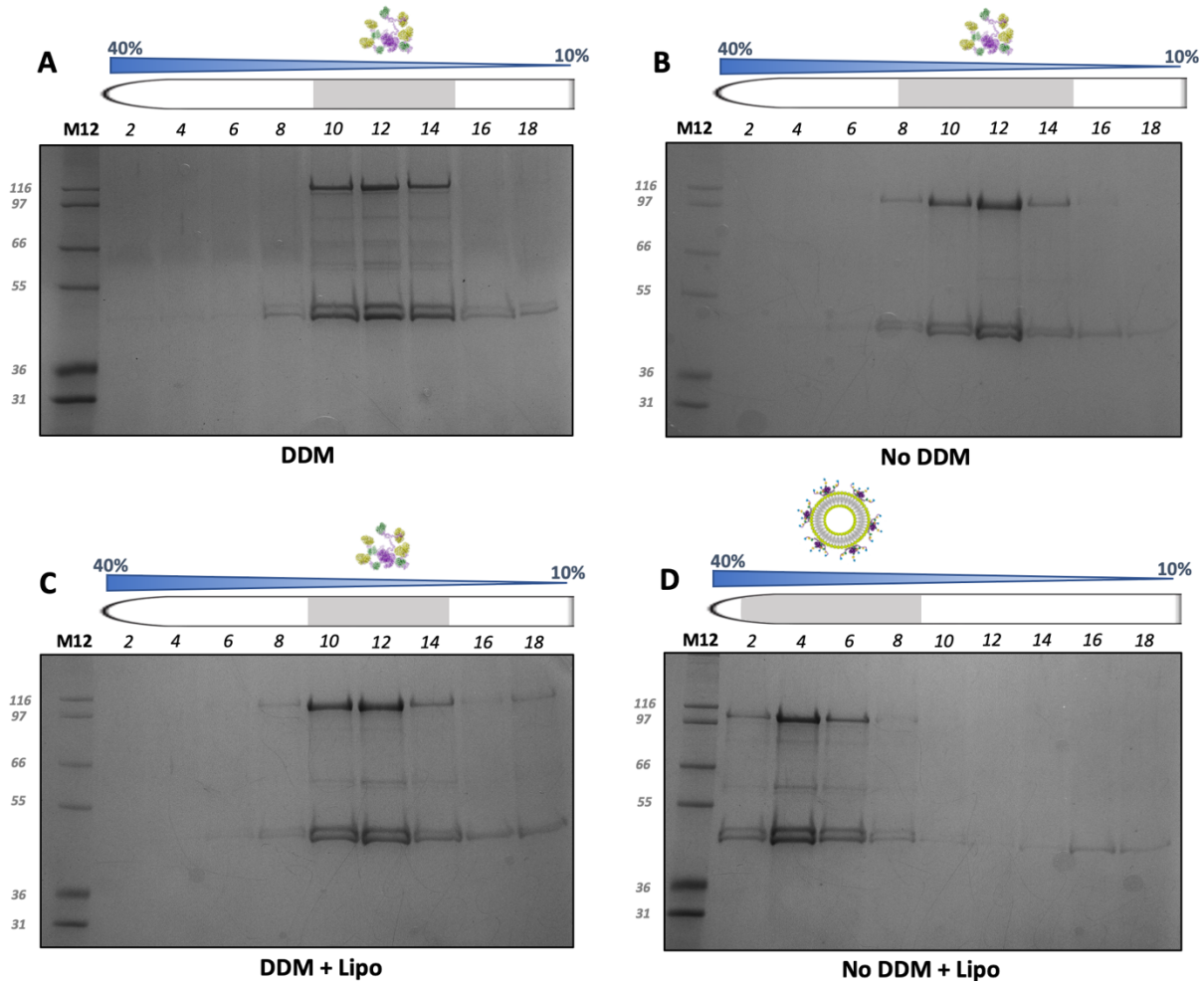


Figure 15: ***Reconstitution of the truncated degradosome on lipid vesicles. (A)*** *In the presence of ß-DDM, the truncated degradosome migrates almost halfway down the glycerol gradient.* ***(B)*** *In the absence of DDM, the truncated degradosome migrates almost halfway down the glycerol gradient too.* ***(C)*** *When both ß-DDM and liposomes are supplemented to the sample, the truncated degradosome does not adhere to the lipid membrane and migrates almost halfway down the glycerol gradient.* ***(D)*** *When the ß-DDM is dialysed out and liposomes are added to the sample, the truncated degradosome binds to the lipid membranes and migrates almost to the bottom of the tube, together with the lipid vesicles. Fractions and glycerol gradients (v/v %) are annotated on top of each gel. Molecular weights of the protein standards are annotated on the left of each gel. Lipo, liposomes; DDM, ß-DDM (0.02% w/v); M12, Mark 12 protein ladder.*

3.3.4.2 Cryo-Electron tomography studies of the membrane-bound truncated degradosome

To further study the membrane-associated truncated degradosome, a cryo-electron tomography (Cryo-ET) approach was used. In short, the truncated degradosome was solubilised in detergent and purified as described in section 3.1.1, then reconstituted on liposomes as described in section 3.3.4.1 and vitrified on cryo-EM grids (Figure

16A). Prior to vitrification, BSA-coated gold fiducials were added to the sample (Materials and Methods 5.11.3.1). Grids were screened on a FEI Talos Arctica (BioCem, Department of Biochemistry, University of Cambridge). Figure 16B shows a representative image of the vitrified liposome-degradosome bodies. Most of the lipid vesicles in the holes were densely packed with the truncated degradosome (Figure 16B, C and E). BSA-coated gold fiducials are dispersed randomly across the ice layer (Figure 16B, yellow arrows). The liposomes are not homogeneous in size, despite extrusion with a 50 nm porous membrane, ranging from 10 nm to 70 nm in diameter. Furthermore, smaller liposomes are more densely packed with protein than larger liposomes (Figure 16C-E). Although most liposomes are saturated with the truncated degradosome, a few liposomes appear to carry distinguishable particles of around 20 nm (Figure 16D). Decreasing the ratio of truncated degradosome:liposome in the sample resulted in more empty liposomes, rather than reducing the overall degradosome packing on the liposomes (data not shown).

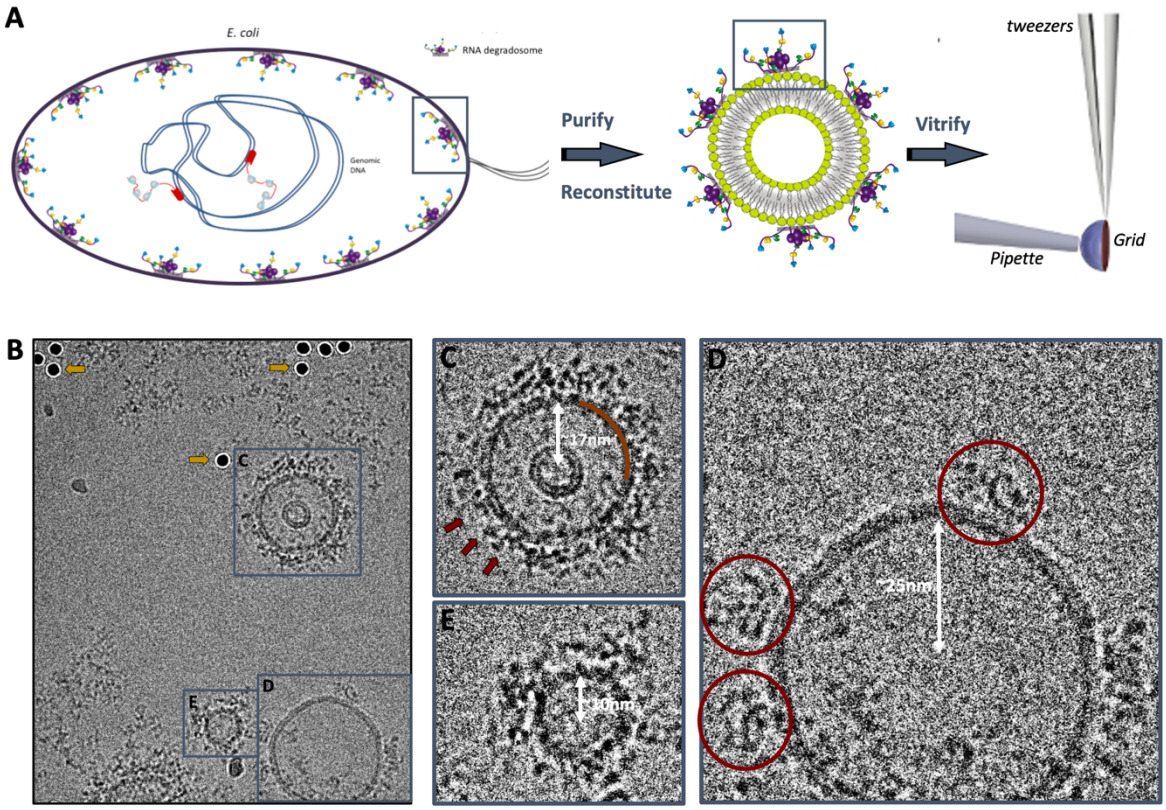


Figure 16: ***Reconstitution, vitrification and imaging of membrane bound truncated degradosome. (A)*** *Schematic overview of the reconstitution procedure. The truncated degradosome was assembled in vivo and solubilised in detergent (ß-DDM) during purification. Next, the detergent was dialysed out and liposomes were added to reconstitute the degradosome on the lipid vesicles, which were then vitrified on cryo-EM grids.* ***(B)*** *A representative cryo-EM image of the degradosome-coated liposomes (blue boxes) (3 µm defocus). Gold fiducials are highlighted with yellow arrows.* ***(C-E)*** *Close ups of selected lipid vesicles. In panel C the lipid membrane is highlighted with a brown curve and densely packed truncated degradosome with red arrows. In panel D individual membrane associated degradosome particles are highlighted with red circles. White arrows and labels show the estimated radius of the lipid vesicles. Schematics in panel (A) were adapted from Schmidly et al., 2018, Dendooven et al., 2017 and Bandyra et al., 2013.*

To further characterise the membrane-bound truncated degradosome, a set of tilt series were collected for a vitrified liposome-degradosome sample on a Talos Arctica (200kV, FEI, BioCem, Department of Biochemistry,

University of Cambridge). The stage was tilted incrementally over a range of 120° (-60° to + 60°, 3° increment) while collecting a low-dose image at every increment (Figure 17A). A bidirectional tilt-scheme was used, going from -20° to 60° and subsequently from -23° to -60°. All tilt series were collected with a Falcon III, in integration mode, at a pixel size of 2.29 Å/pix and 4 µm defocus. Tilt series alignment and weighted back projection (WBP) were performed in Etomo, which is part of the IMOD software package (Kremer *et al.*, 1996) (Materials and Methods section 5.11.3). In short, approximately 20 BSA-coated gold fiducials were tracked throughout each tilt series. This fiducial model was then used to align the individual tilt images globally and locally. The aligned stack was binned twofold and gold fiducials were replaced with random noise. Finally, tomograms were generated via weighted back projection (Figure 17A, right panel). During weighted back projection, high-resolution noise was attenuated by a SIRT-like weighting function to increase contrast (Figure 17B) (SIRT stands for Simultaneous Iterations Reconstruction Technique). In total, six tomograms were reconstructed. From these tomograms a series of spherical and densely packed liposomes were selected for further analysis (Figure 17B and C). Figure 17C shows z-slices through an example of a selected liposome. Ideally the selected liposomes are well-separated and free of neighbouring gold fiducials. Further studies of the selected liposomes, as described below, were carried out in close collaboration with Ms Marta Kubànska, a MPhil student whom I supervised.

Selected liposomes were cropped from the original tomogram in Dynamo (Castaño-Díez *et al.*, 2012) after centring the middle of the lipid vesicle in the sub-tomogram volume. Next, the 1-dimensional radial density profile (1D-RDP) was calculated for each liposome. The 1D-RDP is a radially averaged density profile from the centre of the liposome to the edge of the sub-tomogram volume. Such profiles provide information about the extendedness and potential structural organization of the truncated degradosome when bound to a membrane. To calculate the 1D-RDP, a MATLAB script was written that calculates the 1D-RDP based on spherical symmetrization of the liposome (Materials and Methods 5.11.3). The 1D-RDP script was written together with Mr. Alister Burt (IBS, Grenoble) and relies on a series of symmetry operators from the Dynamo package (Castaño-Díez *et al.*, 2012). The radial density profiles for three densely packed, spherical liposomes are depicted in Figure 17D and an averaged curve is depicted in Figure 17E.

Tracing the density from the middle of the liposome outwards, the first big peak corresponds to the lipid membrane (Figure 17E, grey arrow). Interestingly the density profile reproducibly drops back to the baseline level after the membrane peak. Indeed, for all the liposomes packed with truncated degradosome there is a significant gap in density visible between the membrane and the degradosome. The 1D-RDPs indicate that the width of this gap is approximately 30 Å. The density peak for the truncated degradosome itself is relatively compact and extends 140.5 ($\pm$ 15.7) Å (Figure 17E, horizontal double arrow). This measure of extendedness can be regarded as the real space equivalent of D_{max} in SEC-SAXS (Figure 4). In solution, however, a D_{max} of 427 Å was estimated for the truncated degradosome based on the P(r) function. Due to the membrane-association, the extendedness as measured by the 1D-RDP should be compared to half the D_{max} in solution, i.e. 213.5 Å, because the truncated degradosome cannot traverse through the membrane. Nevertheless, the extendedness of the truncated degradosome bound to a membrane (140.7 Å) is significantly lower than in solution (213.5 Å). This observation

indicates that the truncated degradosome is indeed more compact when associated with a membrane. It should be noted however, that the extendedness estimate from the 1D-RDP is directional. This means that it only represents the extendedness of the truncated degradosome along the normal to the membrane, i.e. away from the membrane. Therefore, an alternative explanation for the discrepancy between the D_{max} in solution and the 1D-RDP is that the truncated degradosome 'spreads' over the membrane like a blooming flower, rather than extending from it like a spike.

Finally, from the individual 1D-RDPs for all selected liposomes, it was apparent that the profile was often bimodal. This was, however, not observed for all the liposomes (Figure 17D and E). Nevertheless, the bimodal shape of the radial density profile could hint towards a higher level of structural organization of the degradosome components on the lipid membrane. This will be further addressed in the next section.

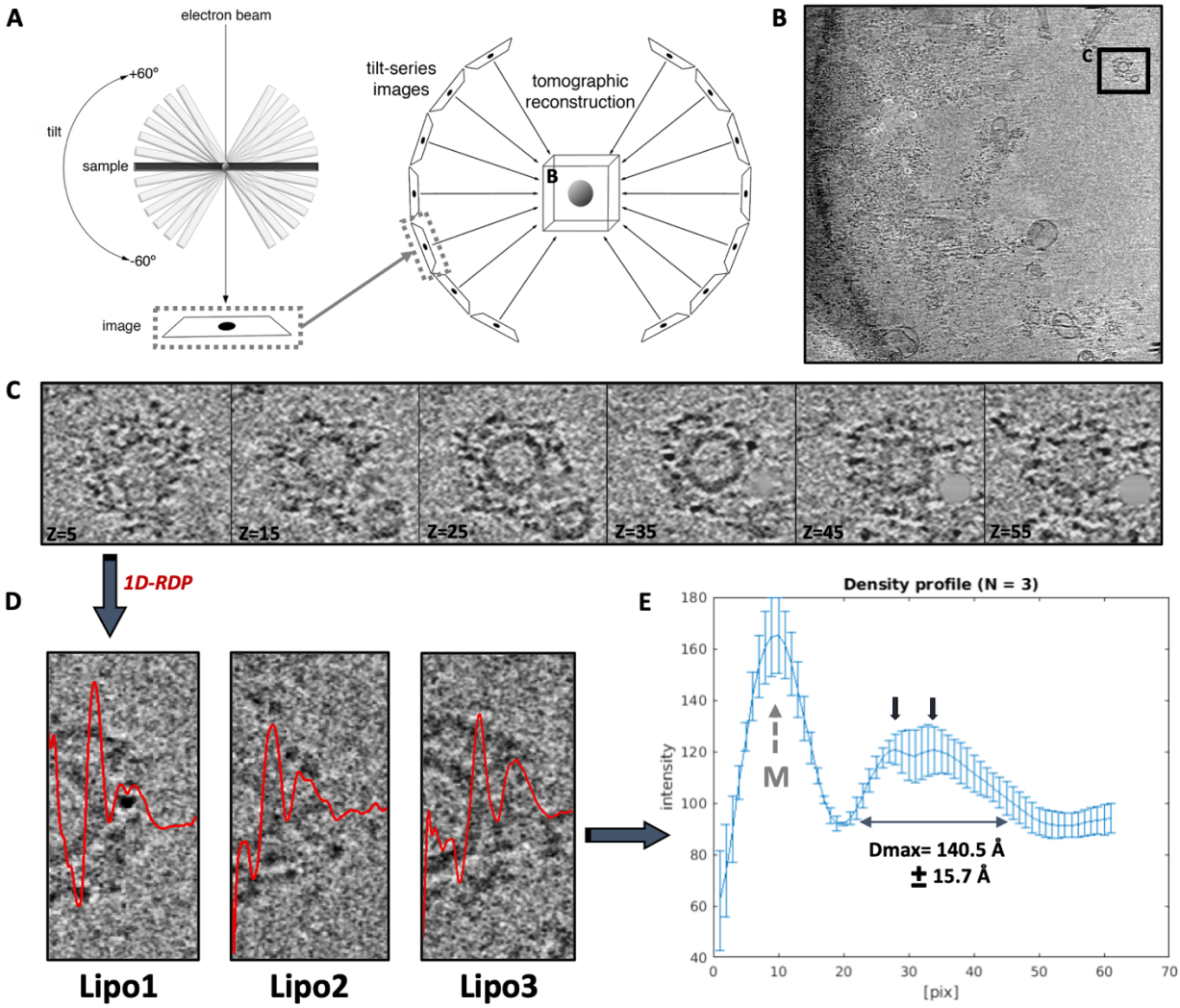


Figure 17: ***Tomogram reconstruction and 1D-RDP calculation for the membrane-bound truncated degradosome****. **(A)** Schematic overview of tilt series collection and tomogram reconstruction. During data collection the stage is tilted incrementally to generate a tilt series of images of the specimen. These tilt series are aligned and used to reconstruct the region of interest in 3D via weighted back projection. **(B)** A central slice through a tomogram of the membrane-bound truncated degradosome. The liposome displayed in (C) is highlighted with a box. The tomogram was filtered with a SIRT-like weighing function. **(C)** Slices through a selected liposome along the Z-axis. Values for Z in each slice correspond to the frame number. The grey sphere at the*

bottom right of the insets corresponds to a gold fiducial that was replaced by random noise. ***(D)*** *1-dimensional radial density profiles (1D-RDP, red) for three different liposomes.* ***(E)*** *Averaged 1D-RDP for the three liposomes depicted in (D). The membrane peak is annotated (M). The extendedness of the membrane bound truncated degradosome was estimated at 140.5 (± 15.7) Å. Black arrows indicate the putative bimodal shape of the radial density profile.*

3.3.4.3 PNPase resides at the periphery of the membrane bound RNA degradosome

As mentioned in the previous paragraph, the components of the membrane bound RNA degradosome might be spatially organised, as indicated by the often bimodal 1D-RDP shape of the truncated degradosome. To further explore this, the full RNA degradosome, i.e. including PNPase, was reconstituted on lipid vesicles as described in section 3.3.4.1 and vitrified for cryo-EM imaging. The aim was to see whether PNPase could be recognised in the EM images. If so, its location in relation to the membrane could further indicate whether there is a degree spatial organisation to the RNA degradosome components. PNPase is a relatively large enzyme (>150kDa) and the cryo-EM studies in chapter II have shown that it aligns with the air-water interface, rendering recognisable, ring-shaped particles in the resulting EM images.

The grids were screened on a FEI Talos Arctica (BioCem, Department of Biochemistry, University of Cambridge) and a representative image is shown in Figure 18A. Many ring-shaped particles, i.e. PNPase, can be recognised in these images. Although they almost exclusively occur in liposome-rich regions on the grid, some PNPase particles are suspiciously far away from a membrane. It is possible that these 'free' PNPase particles have dissociated from the RNA degradosome during grid preparation. Particularly interesting were the lipid vesicles that carried either a complete (Figure 18B) or a partial (Figure 18C) layer of protein. Zooming in on such liposomes revealed that the surrounding PNPase particles are further away from the membrane than the measured extendedness of the truncated degradosome (Figure 18B and C, light blue line). It should be noted that PNPase binds RNase E at the very end of its scaffold domain, and that the full-length RNase E has a scaffold domain that is more than 200 residues longer than that of the truncated degradosome. The apparent location of PNPase at the periphery of the membrane-bound protein density further points towards a spatial organisation of the RNA degradosome when bound to a membrane, assuming that PNPase is indeed still associated to RNase E in these samples.

It is clear that the association of PNPase with the air-water interface can disturb the putative degradosome architecture significantly, which would question the interpretation of the observations described above. If PNPase can be kept away from the air-water interface during grid freezing, then tilt series could be collected for these images to more accurately pinpoint its location with respect to the lipid membrane. In chapter II it was found that the zwitter-ionic detergent CHAPSO keeps PNPase away from the air-water interface. Unfortunately, adding CHAPSO to the liposome samples, even below the critical micelle concentration (CMC) resulted in distorted liposomes, devoid of bound degradosome (data not shown).

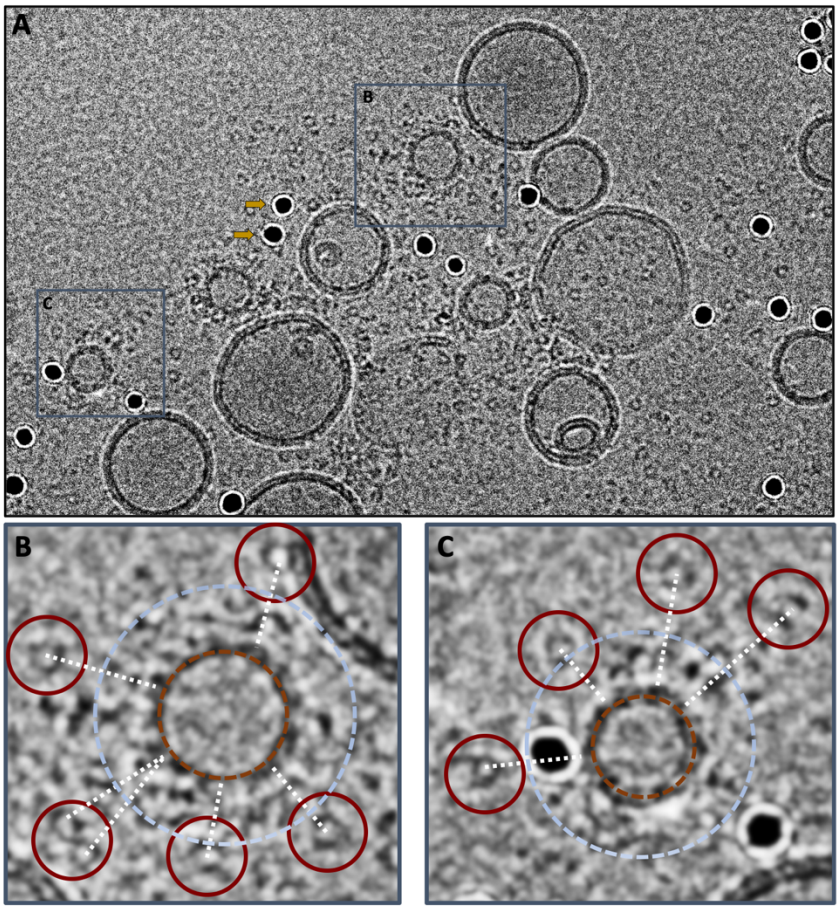


Figure 18: ***Cryo-EM studies of the membrane-bound full-length RNA degradosome. (A)*** *Representative image of the liposome-bound full-length RNA degradosome (3 µm defocus). Yellow arrows point to two gold fiducials* ***(B-C)*** *PNPase (red circles) is located further away from the lipid membrane (brown circle) than the average extendedness of the truncated degradosome (light blue circle).*

3.3.5 Cryo-ET studies of the 30S small ribosomal subunit bound to the membrane-associated truncated degradosome

The ultimate goal is to resolve the membrane bound truncated degradosome, or the full RNA degradosome, by aligning and averaging individually picked degradosome volumes on the lipid membrane, a process called sub-tomogram averaging. The liposome saturation with truncated degradosome complicates defining individual truncated degradosome volumes for further processing by sub-tomogram averaging. In section 3.3.3, interactions between the 30S small ribosomal subunit and the recognition core/truncated degradosome were studied. It was found that although both the recognition core and the truncated degradosome bind the 30S small ribosomal subunit to form putative RNA surveillance assemblies, they remain conformationally heterogeneous when doing so. In this section, the interactions between the 30S small ribosomal subunit and the membrane bound truncated degradosome are studied with ultracentrifugation assays and cryo-electron tomography. Since the RNase E scaffold domains interact with both lipid membranes and ribosomes or ribosomal subunits, the aim was to visualise and study super assemblies that combine these interactions, i.e. liposome-bound truncated degradosome-30S assemblies.

3.3.5.1 The 70S ribosome and 30S small ribosomal subunit bind the membrane bound truncated degradosome

To test whether the 30S small ribosomal subunit and the 70S ribosome bind the membrane-bound truncated degradosome, the ultracentrifugation assay described in section 3.3.4.1 was used. These ultracentrifugation assays were carried out in close collaboration with Ms. Giulia Paris, a master student whom I supervised. The truncated degradosome was reconstituted on liposomes first and 30S or 70S was added for the last two hours of the dialysis step. Two negative controls were set up in parallel: 30S alone and 30S supplemented with liposomes. During the ultracentrifugation step, the 30S migrated about halfway down the glycerol gradient, both without and with liposomes present in the sample (Figure 19A and B). These observations show that the 30S ribosome does not bind lipid membranes on its own. However, when the truncated degradosome is reconstituted on the lipid membrane prior to adding the 30S subunit or the 70S ribosome, there is a significant enrichment of the latter two near the bottom of the centrifugation gradient (Figure 19C and D). In conclusion, both the 30S small ribosomal subunit and the whole 70S ribosome associate with the membrane-bound truncated degradosome and co-migrate down the glycerol gradient.

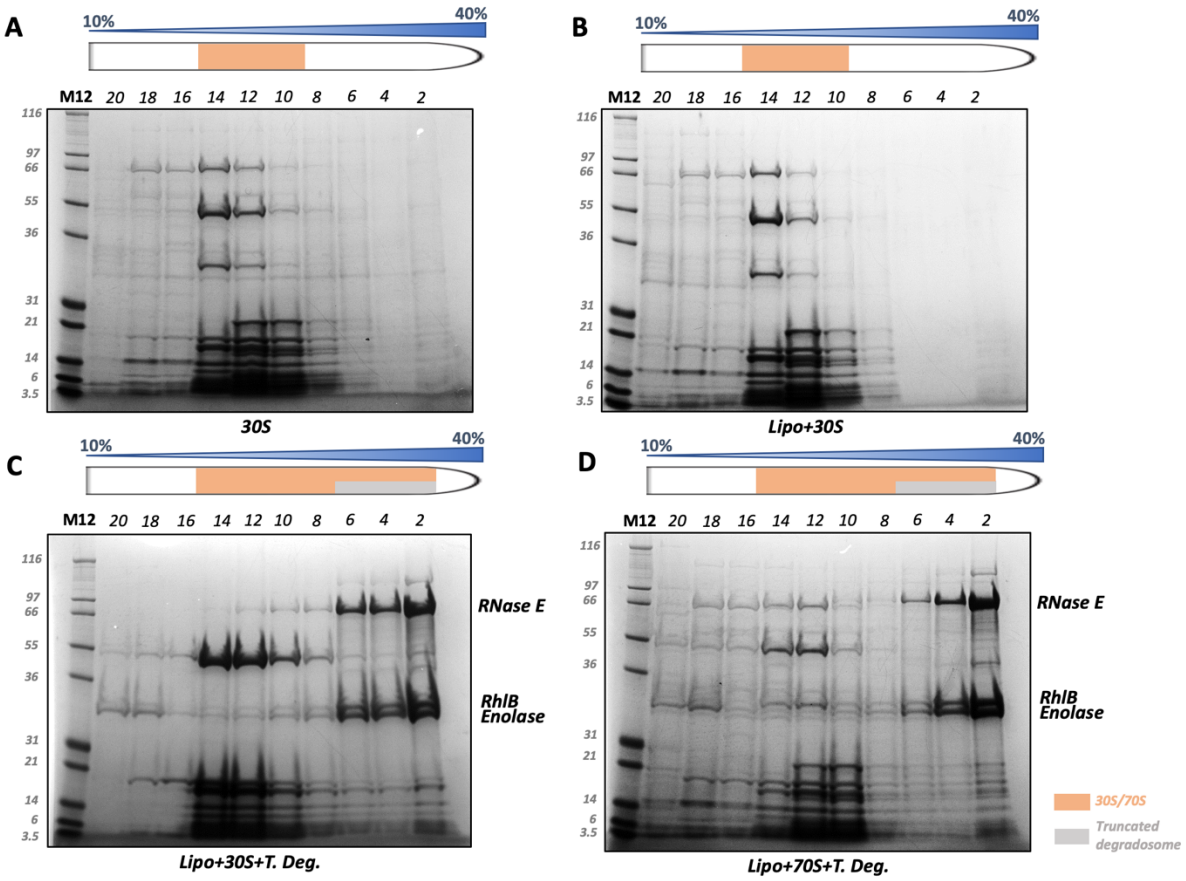


Figure 19: ***Ultracentrifugation assays of the 70S ribosome and 30S small ribosomal subunit (A)*** *The 30S small ribosomal subunit migrates halfway down the glycerol gradient during the ultracentrifugation step.* ***(B)*** *The 30S small ribosomal subunit does not adhere to lipid membranes and migrates halfway down the glycerol gradient when liposomes are present in the sample.* ***(C)*** *When the truncated degradosome is bound to the lipid vesicles, a fraction of the 30S small ribosomal subunit co-migrates down the glycerol gradient with the liposomes.* ***(D)*** *Same as (C) but for the 70S ribosome. Fractions and glycerol gradients (v/v %) are annotated on top of each gel. Molecular weights of the protein standards are annotated on the left of each gel. Lipo, liposomes; T. Deg., truncated degradosome; M12, Mark 12 protein ladder.*

3.3.5.2 Sub-tomogram averaging and localisation of 30S on lipid vesicles

To further study the interaction between the 30S small ribosomal subunit and the membrane-bound truncated degradosome, Cryo-ET studies were carried out. The 30S-truncated degradosome-liposome sample was supplemented with BSA-coated gold fiducials and vitrified on cryo-EM grids with a Vitrobot Mark IV, FEI, BioCem, Department of Biochemistry, University of Cambridge). Grids were screened with a Talos Arctica (200kV, FEI, BioCem, Department of Biochemistry, University of Cambridge). A representative image is depicted in Figure 20A. When zooming in on the lipid vesicles in the image, it appears that a subset of the 30S particles are bound to the truncated degradosome, which is in turn associated with the vesicle membrane (Figure 20B and C, blue arrows). It should be noted that the images in Figure 20A and B are integrated along the viewing axis, which means that features that are on top of each other in the vitreous ice will overlap in the images. To better separate 30S small ribosomal subunits that are associated with the truncated degradosome, tilt series were collected with the same settings as described in section 3.3.4.2. The tilt series images were aligned in IMOD (Kremer *et al.*, 1996) and

tomograms were generated from the aligned image stack (binned 4x) via WBP. For visualisation, a SIRT-like weighting function was used during WBP to minimise high resolution noise and features. These tomograms revealed that free 30S ribosomal subunits align with the air-water interfaces at the top and bottom of the vitreous ice (Figure 20F, bottom panel). The 30S subunits that are associated with the truncated degradosome, however, do not reside at the air-water interface, and form semi-discrete particles near the liposome membranes (Figure 20D, blue arrows).

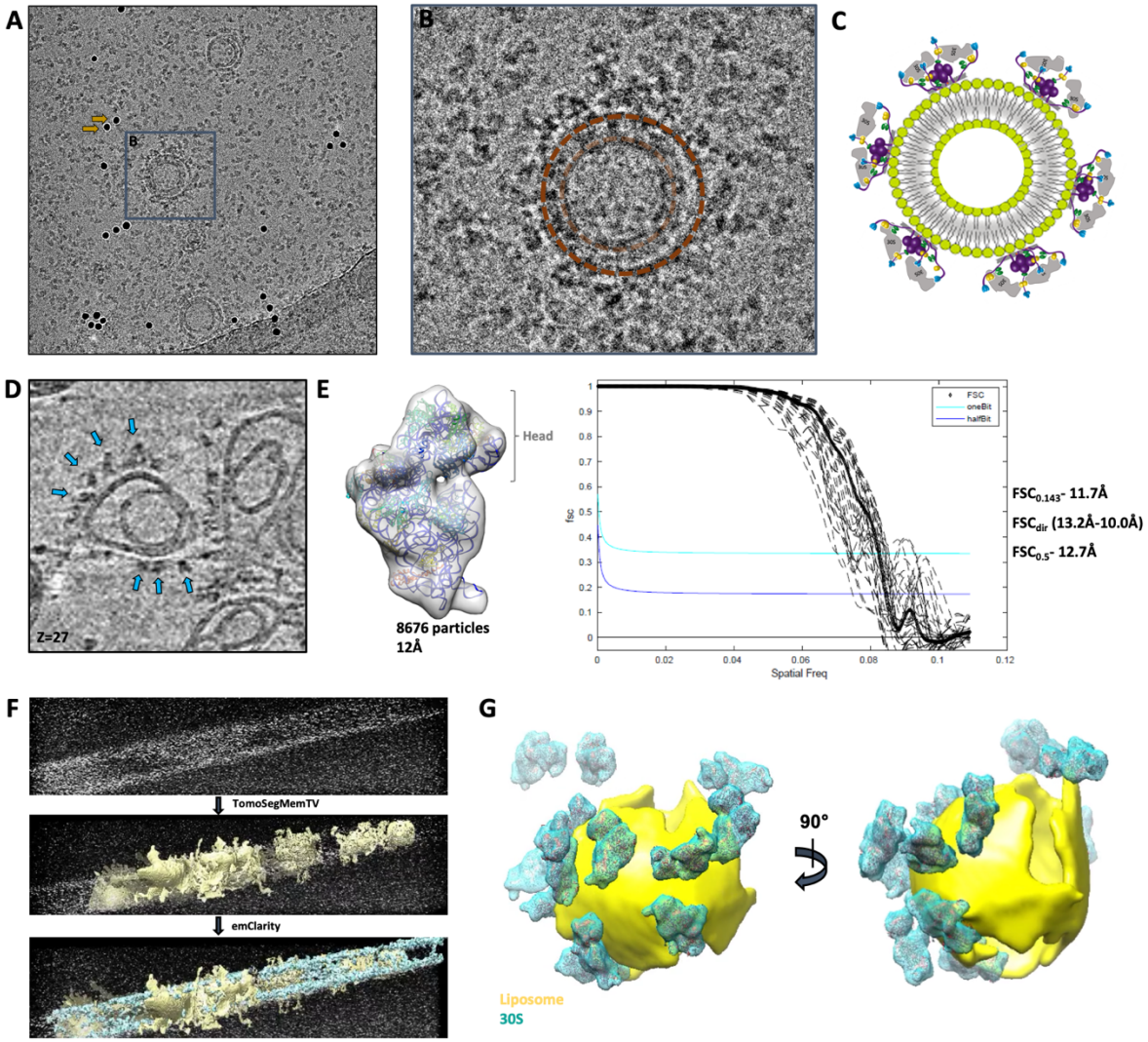


Figure 20: ***Tomographic studies of the 30S small ribosomal subunit bound to the membrane-associated truncated degradosome****. **(A)** Representative cryo-EM image of truncated degradosome-coated liposomes and 30S small ribosomal subunits (3 µm defocus). Two gold fiducials are highlighted by yellow arrows. **(B)** Magnified representation of a liposome (modelled with a dark brown circle), which coincidentally encloses a second liposome (light brown circle). 30S particles are enriched on the outer liposome, which is saturated with truncated degradosome. **(C)** Schematic representation of a lipid vesicle coated with degradosome-30S super-complexes. **(D)** Magnified image of a slice through a reconstructed tomogram showing distinct 30S particles near the lipid membrane of a liposome (blue arrows). **(E)** Sub-tomogram average as reconstructed with emClarity. 8676 sub-tomogram volumes contributed to this average, which has an estimated resolution of 11.7 Å (GS-FSC, right panel). The dotted lines in the FSC plot correspond to directional FSC curves. The directional resolution estimates range from 10 Å to 13.2 Å. **(F)** The raw, noisy tomograms (top) were processed with TomoSegMemTV to model the liposome membranes*

(middle, membranes are modelled in yellow). A spatial distribution volume of the 30S particles was generated by mapping back the reconstructed averages to replace the original, noisy 30S particles in the tomogram (bottom, 30S averages are in cyan). From this distribution it is apparent that the bulk of the 30S particles align with the two air-water interfaces. ***(G)*** *Magnified image of a segment of the distribution map in (F), showing a modelled liposome (yellow) and a series of 30S small ribosomal subunits (cyan) in close proximity.*

Due to severe missing wedge effects, arising from the limited range of tilt angles during data collection, it is hard to recognise detailed features in a raw tomogram. To resolve and further study the 30S small ribosomal subunit particles, emClarity was used for sub-tomogram averaging (Himes *et al.*, 2018). The goal was to see if additional density for the truncated degradosome and the lipid membrane would appear in the 30S 3D reconstruction. If the latter were the case, then the truncated degradosome would have been resolved '*de novo*', i.e. without a template, by aligning the 30S sub-tomogram volumes. To this end, all ribosomes, both bound to the truncated degradosome and floating freely in the vitreous ice, were picked via automated template matching in emClarity and inspected manually. As a reference for template matching, a map was generated from pdbID 4ADV in UCSF Chimera (Pettersen *et al.*, 2004) and lowpass filtered to 30 Å. An initial reference for subsequent alignments was generated from a picked particle set of 8676 30S sub-tomogram volumes. The tilt series alignments were refined via two rounds of TomoCPR. TomoCPR is a tool within emClarity that refines the tilt series alignment by using individual sub-tomogram volumes, i.e. 30S particles, as fiducial markers. After several rounds of aligning and averaging the 30S sub-tomogram volumes a consensus map of the 30S ribosome was reconstructed at 11.7 Å resolution (GS-FSC) (Figure 20E). The FSC-curve in Figure 20E indicates that the reconstruction is isotropic and the dotted FSC curves, i.e. directional resolution estimates, correspond to a range of 10-13.2 Å. Given that these data were collected on a 200kV microscope with a Falcon III camera (integration mode), the achieved global resolution is reasonable.

Unfortunately, no additional density could be observed for either the truncated degradosome or a lipid membrane in the sub-tomographic averages. A likely explanation is that the majority (>90%) of the 30S particles that contributed to the final average are not bound to the truncated degradosome. If so, additional densities will be averaged out during the alignments due to low occupancy. To better visualise the putative 30S particles associated with the truncated degradosome in the raw tomograms, the reconstructed 30S average was mapped back in the original tomogram to replace the noisy sub-tomogram volumes that contributed to it, while conserving their position and orientation (Figure 20F, cyan). This resulted in an interpretable spatial distribution map of the 30S particles within the original tomogram and indicated that the majority of 30S particles indeed align with the air-water interfaces, i.e. not bound to the proteoliposomes (Figure 20F). However, a small subset of the 30S volumes are located away from the air-water interface and in close proximity to the lipid membranes, which were traced and modelled with TomoSegMemTV (Martinez-Sanchez *et al.*, 2014) (Figure 20F and G, yellow). Interestingly, 30S particles near liposomes appear to reproducibly face the lipid membrane along the longest particle axis, which could indicate the interaction between the truncated degradosome and the 30S small ribosomal subunit is specific. The next step in the analysis is to separate free 30S particles from those bound to proteoliposomes via 3D classification approaches. Aligning and averaging the latter might resolve the membrane-bound truncated

degradosome. Unfortunately, attempts to do this via principle component analysis with emClarity have been unsuccessful so far.

.

4 Discussion

Over the past 25 years, many studies have elucidated functional roles of bacterial RNA degradosome assemblies in post transcriptional gene regulation in the cell. In addition, many degradosome components have been elucidated structurally, which has contributed important insights in RNA degradation mechanics. However, structural studies of an RNA degradosome in its entirety and methods to do so have been elusive for many years. For one, the conserved intrinsic flexibility of the RNase E C-terminal half in Gram-negative bacteria, which serves as a scaffold domain for other degradosome enzymes, impedes structural studies of the assembly. Therefore, the term 'un-structural biology' was coined to summarise the efforts to explore the layout of the RNA degradosome. In *E. coli*, the RNA degradosome is tethered to the bacterial inner membrane via an amphipathic helix on the RNase E scaffold domain, which further complicates sample preparation and subsequent setup of adequate experiments. Both these features, intrinsic disorder of the RNase E scaffold domain and cellular localisation, are inherent to the RNA degradosome and have been addressed in terms of their putative roles in efficient riboregulation in the cell (see Introduction). In this chapter, approaches to tackle these conserved degradosome characteristics were explored.

4.1 Novel purification approaches for the RNA degradosome

A reproducible protocol was developed for the expression and purification of both the entire RNA degradosome and the truncated degradosome, a large sub-assembly of the RNA degradosome (section 3.1). From the results described in section 3.1, it is clear that adequate inhibition of endogenous proteases during early purification steps and solubilisation of the truncated degradosome with Triton-x 100 and ß-DDM were two key improvements to existing strategies for RNA degradosome reconstitution and purification (Worral *et al.*, 2008). In addition, using high salt buffers early on in the purification and adding a cation exchange purification step to the protocol ensures negligible levels of RNA contamination in the purified sample. Including reducing agents such as TCEP is also an important additive to ensure stability of the degradosome. Taken together, this new and highly reproducible reconstitution protocol for the truncated and full RNA degradosome results in intact protein assemblies of high purity. Moreover, the high protein yields achieved with this protocol were sufficient for downstream structural and biophysical studies that require high protein concentrations, e.g. SEC-SAXS.

4.2 Biophysical approaches reveal the conformational behaviour of the RNA degradosome

To test whether the *in vivo* reconstituted truncated degradosome has the expected oligomeric state, a series of biophysical assays were carried out. An analytical ultracentrifugation (AUC) experiment showed that the purified truncated degradosome is tetrameric, with an estimated molar mass of >800 kDa (Figure 3A). A similar outcome was observed for dynamic light scattering experiments, where truncated degradosome was predicted to be a homogeneous assembly in solution, with a hydrodynamic size of 330Å (Figure 3B). AUC and DLS experiments also indicate that the assembly is likely to be extended. This is in line with recent experiments studying the conformational heterogeneity of a fragment of the RNase E scaffold domain (Bruce *et al.*, 2018). The construct

studied here consists of RNase E (603-850), with two RNA binding sites, Enolase and RhlB, together forming the so called 'recognition core'. Bruce *et al.* found that association of Enolase and RhlB with the RNase E scaffold domain might reduce its conformational heterogeneity. To probe the conformational heterogeneity of the RNA degradosome, SEC-SAXS experiments were carried out on the truncated degradosome, which harbours four of these recognition cores (one for each RNase E scaffold domain). The large Radius of gyration R_g and maximum interatomic distance D_{max} and the normalised Kratky plot, with its characteristic large secondary 'bump', are signatures for the conformational heterogeneity of the truncated degradosome (Figure 4). In fact, these observations indicate that beyond its structured domains and binding partners, the RNA degradosome as a whole is 'unstructured' in solution. Interestingly, when 9S rRNA, a known substrate of RNase E, binds the truncated degradosome, the latter forms a more compact assembly (Figure 5). It is likely that 9S rRNA binds the RNA binding sites on the RNase E scaffold domain and the DEAD-box helicase RhlB, in addition to the catalytic core of RNase E. These interactions with a substrate RNA could fix the truncated degradosome in a more rigid state.

The observation that the truncated degradosome is an assembly with little rigidity complicates further structural studies. However, the RNase E scaffold domain is the only part of the RNA degradosome that has not been modelled yet based on structural data. Good approaches to model the conformational behaviour of something as flexible as the RNase E scaffold domain are hard to design. An established approach is that of ensemble optimisation modelling (EOM), an integrative structural biology method that aims to explore the conformational landscape of a protein based on prior knowledge and experimental scattering profiles (e.g. SEC-SAXS) (Bernado *et al.*, 2007; Tria *et al.*, 2015). A schematic overview of the EOM strategy is given in (Figure 6). In short, EOM scans a pool of random structures of the protein of interest to find ensembles of models that together explain the experimental data. A large pool of random structures of the truncated degradosome was generated with a series of scripts developed together with Dr. Aliaksandr Shkumatau (Figure 7). The random truncated degradosome pool samples conformational space for the four RNase E scaffold domains while incorporating *a priori* information about the structured domains (Koslover *et al.*, 2008; Nurmohamed *et al.*, 2009; Bruce *et al.*, 2018; Chandran *et al.*, 2007). Ensembles of 4-7 truncated degradosome models were selected from the random pool in every round (for 100 rounds), based on the experimental SAXS data, and their structural characteristics recorded (Figure 8A and B). From these analyses it is clear that the truncated degradosome adopts highly extended conformations in solution, with very little rigidity (R_{flex}=79.01%) (Figure 8D). It should be noted that the fit of the ensembles against the experimental scattering data is not optimal at low q ranges. Given the size distribution functions in Figure 8 it is likely that extended structures were underrepresented in the random pool, leading to under-sampling of such regions of conformational space. Current efforts are aiming to shift the average size of the models in the random pool to include more extended candidate structures to complement the current results. In conclusion, the SAXS experiments of the tetrameric truncated degradosome provide the first ever description of its conformational behaviour in solution.

4.3 Cryo-EM approaches to visualize the solubilised truncated degradosome

The outcomes of the SEC-SAXS experiments suggest that attempts to solve a structure for the truncated degradosome might be futile. Interestingly, the 30S small ribosomal subunit and the 70S ribosome have been shown to interact with the RNA degradosome *in vitro* (Tsai *et al.*, 2012) and both are predicted to form RNA surveillance assemblies *in vivo* that help assess the 'viability' of a transcript before or during translation (Dendooven *et al.*, 2020). Such RNA surveillance assemblies are reasonably well studied in eukaryotes, where ribosomes have been shown to interact with the exosome, which is the Eukaryotic equivalent of the bacterial RNA degradosome, during translation (Schmidt *et al.*, 2016; Dendooven *et al.*, 2020). Tsai *et al.* (2012) found that mainly the degradosome recognition core interacts with ribosomes in bacteria. Interactions between one or more 30S subunits and the scaffold domains of the truncated RNA degradosome could result in a more rigid super complex. Cryo-EM studies of the 'recognition core', bound to the 30S small ribosomal subunit reveal weak cryo-EM density near the head of 30S after many rounds of careful 3D classification (Figure 14A and B). Unfortunately, conformational heterogeneity of the recognition core did not allow for high resolution reconstructions or rigid body docking of existing models of the recognition core components. However, this low-resolution reconstruction is the first visualisation of the RNA degradosome scaffold region and its participation in a putative surveillance assembly.

Besides high levels of conformational heterogeneity, it is possible that association of the truncated degradosome with the air-water interface (AWI) during grid preparation results in partial or complete denaturation of the assembly, which in turn complicates 3D reconstructions (Noble *et al.*, 2018; D'Imprima *et al.*, 2019). To explore these potential issues, cryo-EM studies were carried out on the RNase E catalytic core (RNase E 1-529). Even though the RNase E N-terminal half is not the focus of this chapter, the idea was that successful cryo-EM reconstructions of this catalytic tetramer would help pinpoint problematic regions of the RNA degradosome for cryo-EM studies by elimination. The N-terminal domain of RNase E has been crystallised several times and was found to adopt different conformational states upon binding RNA substrates (Callaghan *et al.*, 2005, Koslover *et al.*, 2008, Bandyra *et al.*, 2018). As for PNPase in Chapter II, adding the zwitter ionic detergent CHAPSO to the cryo-EM sample was necessary to keep the RNase E catalytic core away from the AWI during grid freezing (Figure 11A). It should be noted that both PNPase and RNase E contain S1 domains, just like the 30S small ribosomal subunit, and that CHAPSO prevents association with the air-water interface for all three species. 3D reconstructions of the RNase E catalytic core resulted in an interpretable consensus reconstruction (Figure 11B-D). Subsequent 3D variability analyses in cryoSPARC resolved three modes of conformational heterogeneity within the catalytic core, corresponding to a 'rocking', 'rolling' and 'breathing' motion of the RNase E protomers (Figure 12B) (Punjani *et al.*, 2020). It is likely that this inherent flexibility in the RNase E core adds to the overall flexibility of the RNA degradosome, albeit to a relatively limited extent. In conclusion, the studies described in this paragraph show that the main culprit in calculating 3D reconstructions of the truncated degradosome particles with the currently available tools are the RNase E scaffold domains.

4.4 Cryo-ET approaches to visualize the membrane-bound truncated degradosome

A protocol was developed to reconstitute the truncated degradosome on lipid vesicles (Figure 15). Reconstitution on a lipid membrane not only renders the degradosome in a more native environment, free of detergent, it might also fix the RNase E scaffold domains in a more rigid conformation. Moreover, the lipid vesicles could keep the truncated degradosome away from the AWI during grid freezing. Cryo-EM grids were prepared for these proteo-liposomes and tilt series collected for tomographic reconstructions. Although many liposomes were coated with truncated degradosome, most of the liposomes were too densely packed for further sub-tomogram alignments and averaging (Figure 16).

To explore the extendedness of the truncated degradosome when bound to lipid membranes, 1-dimensional radial density profiles (1D-RDP) were calculated for several liposomes (Figure 17). These density profiles revealed a reproducible gap between the lipid membrane and the truncated degradosome density which suggests that the main body of the truncated degradosome is separated from the membrane (Figure 17D and E). In addition, the extendedness of the truncated degradosome on the membrane was estimated to be 140.5 (± 15.7) Å along the normal of the membrane. This is much more compact than the truncated degradosome in solution, estimated at 427 Å by SEC-SAXS experiments or 213.5 Å when taking into account the membrane barrier. It is possible that the truncated degradosome mainly extends along the membrane rather than perpendicular to it, which could explain the discrepancy in observed extendedness for the soluble and membrane-bound truncated degradosomes. The SEC-SAXS results described in section 3.2.3.2 indicate that the truncated degradosome compacts upon binding RNA substrates, like 9S rRNA. Further tomography studies of membrane-bound truncated degradosome engaged on 9S rRNA have been initiated and will elucidate whether interactions with the lipid membrane and 9S rRNA have a cumulative effect on assembly extendedness.

The RNA degradosome works as an integrated molecular machine in the cell, digesting RNA substrates down to individual ribonucleosides. Some degree of organization encoded within its inherent conformational chaos is likely to be necessary to facilitate the processing order of a transcript and ensure completion of the RNA degradation pipeline within a single assembly. Interestingly, the averaged 1D-RDPs appear to be bimodal, which hints towards spatial organization of the truncated degradosome components with respect to the membrane (Figure 17E, black arrows). This observation, however, was not entirely reproducible across the selected liposomes and needs to be further investigated in future tomograms. Interestingly, cryo-EM images with the whole RNA degradosome reconstituted on liposomes suggest that PNPase mainly resides at the periphery of the membrane-bound protein density (Figure 18). This too points towards a degree of order in the organization of degradosome components in relation to the membrane.

4.5 An inclusive Cryo-ET approach to visualize membrane-bound surveillance assemblies

The experiments so far indicate that the RNA degradosome might adopt a more defined conformation upon binding RNA substrates or associating with a lipid membrane. Both these interactions are at least partly established at the RNase E scaffold domain, which embodies the inherent conformational heterogeneity of the truncated degradosome. As described in section 4.3, the 30S small ribosomal subunit also associates with the RNase E scaffold domain to form putative RNA surveillance assemblies, which could be conformationally more rigid. In an attempt to bring these findings together, cryo-ET studies were carried out on the 30S small ribosomal subunit in complex with the membrane bound truncated degradosome (section 3.3.5).

The long-term goal of this approach is to resolve the membrane-bound truncated degradosome 'de novo' via sub-tomogram alignments and -averaging of the 30S particles that are bound to it. The 30S small ribosomal subunit contains many electron-rich phosphate molecules in its RNA content and therefore contributes a more signal to the cryo-EM image than proteinogenic particles, which helps with particle alignments. To this end, all 30S volumes were extracted from 30 tomograms and aligned/averaged with emClarity (Himes *et al.*, 2018). These sub-tomogram averaging efforts and subsequent mapping of the averages in the original tomogram revealed that a subset of the 30S volumes are likely to be bound to the membrane-associated truncated degradosome (Figure 20E). Given that more than 90% of all 30S particles are not associated with the membrane-bound truncated degradosome, no additional density for the truncated degradosome or the lipid membrane was resolved. Unfortunately, 3D classification of sub-tomogram volumes in emClarity is not optimal for resolving large entities of extra density such as the truncated degradosome or lipid membranes and other classification tools are being explored. Alternatively, particle picking could be restricted to the membrane bound super assemblies in the tomogram. Several approaches for picking membrane bound proteins in tomograms have been developed in PEET and Dynamo and are currently being explored (Castaño-Díez *et al.*, 2017, Nicastro *et al.*, 2006; Heumann *et al.*, 2011).

The interactions between the 30S small ribosomal subunit and the RNA degradosome have been mostly described as an experimental trick to help resolve the RNA degradosome scaffold domain. However, in addition to interactions with the 70S ribosome (Tsai *et al.*, 2012), it is possible that the degradosome interacts with 30S-based translation initiation assemblies such as the 30S pre-initiation complex (PIC) *in vivo* as a fidelity check (Hussain *et al.*, 2016). The latter is an assembly in which three initiation factors help pre-organize the start codon of a mRNA at the P-site of the 30S ribosomal subunit and enable selecting initiator tRNA before full ribosome assembly and translation elongation (Milón & Rodnina, 2012). If the transcript engaged by the PIC is aberrant, the RNA degradosome would receive a signal, e.g. from a small non-coding RNA (sRNA) targeting the faulty mRNA, and cleave the substrate. This hypothesis is currently being investigated in the Luisi group, and if true, the structural elucidation of a membrane-bound degradosome-30S super complex would significantly add to understanding this process.

4.6 Towards an accurate structural model of the bacterial RNA degradosome

The approaches described in this Chapter raise the question as to what the ultimate strategy is to structurally elucidate the conformational behaviour and the working mechanism of the largely unstructured RNA degradosome. The most promising approach is undoubtedly that of cryo-ET and sub-tomogram averaging. Of particular interest would be the exploration of the potential spatial organisation of the RNA degradosome components when the assembly is bound to a lipid membrane. An interesting approach to further study this would be to purify various sub-assemblies of the RNA degradosome and reconstitute them on lipid vesicles for tomographic studies (Figure 21). Differences in the calculated 1D-RDP profiles for each degradosome sub-assembly can then be monitored and could further hint towards a degree of loose quaternary order of the RNA degradosome. Since grid-preparation, data collection/processing pipelines and 1D-RDP calculation protocols have been established in this chapter, the biggest bottleneck will be sample preparation. For example, preliminary attempts to purify RNase (1-1061) without any of the other degradosome components have been unsuccessful (data not shown). Although the protein yields of full-length RNase E are high, the other degradosome components co-purify endogenously. Specific deletion-strains will have to be tested to overcome unwanted co-purification.

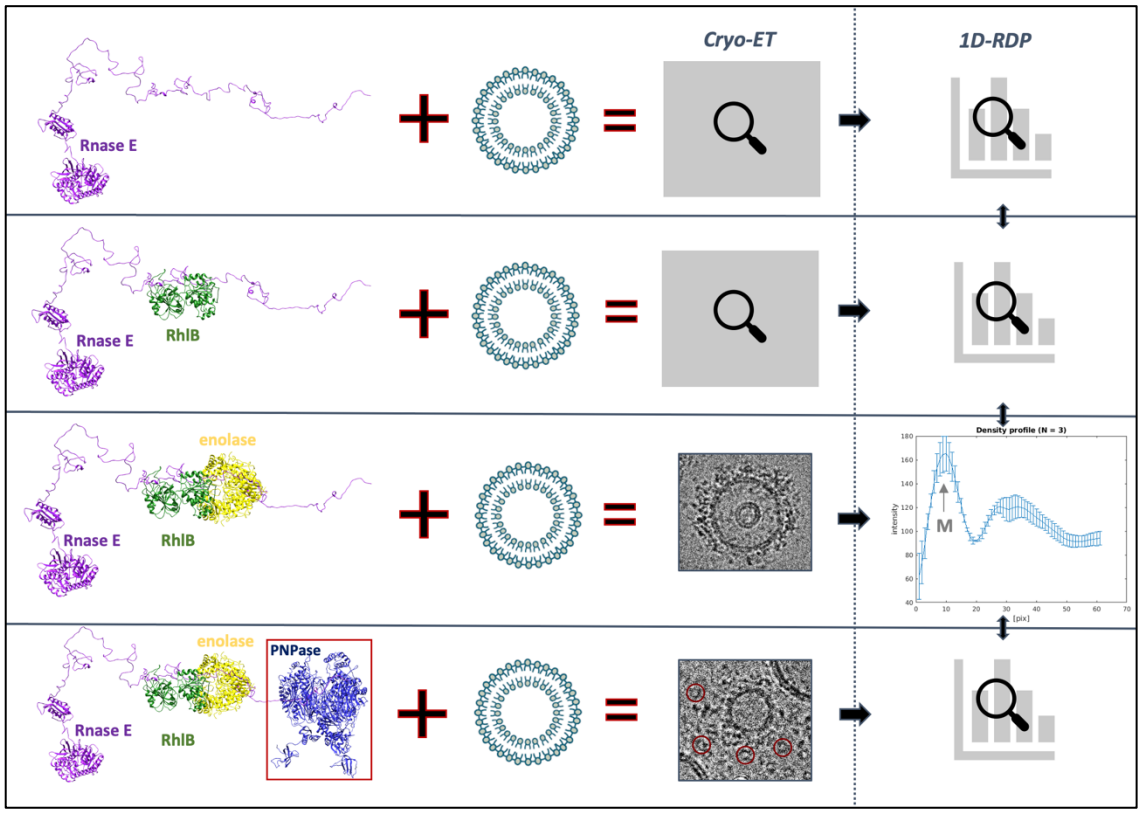


Figure 21: ***Schematic presentation of a proposed experimental strategy to study the spatial organisation of the RNA degradosome components****. Different RNA degradosome subassemblies can be purified and reconstituted on lipid vesicles prior to cryo-ET studies. 1D-RDPs can be calculated from tomographic volumes of protein coated liposomes and compared for different subassemblies. For example, PNPase appears to mainly reside at the periphery of the membrane-bound RNA degradosome (bottom inset, red square and circles).*

Given the extreme conformational heterogeneity of the RNA degradosome, a dedicated sub-tomogram averaging and classification scheme can be designed in the future. In particular, tomographic volumes of the RNA

degradosome could be dissected *in silico* into its individual components via careful particle picking and extensive 3D classification procedures. Accurate particle recognition in tomograms and 3D classification are perhaps the most interesting challenges in the field of cryo-STAC today (Sub-Tomogram Averaging and Classification). Different strategies to tackle this challenge are being developed (e.g. Martinez-Sanchez *et al.*, 2020). For the degradosome, a tool that performs multi-template matching with spatial constraints (i.e. close to a lipid membrane) would be a significant step forward. Once all the individual degradosome partners are aligned and averaged separately, they can be mapped back into the original tomogram to compose the membrane-bound RNA degradosomes and examine their layout. The two ribonucleases of the RNA degradosome, RNase E and PNPase, each have a molar mass of more than 180 kDa, which increases the success rate of identifying these in a tomogram and aligning the corresponding sub-tomogram volumes. Enolase and RhlB, however, are much smaller (less than 100 kDa), which would make both particle-picking and alignments nearly impossible with the current cryo-EM hardware (electron microscopes/detectors) and software (picking/reconstruction algorithms). Luckily the cryo-ET field is developing rapidly, and the strategy proposed in this last paragraph will soon be testable, hopefully leading to the first ever structure of the unstructured bacterial RNA degradosome.

5 Materials and Methods

5.1 RNase E mutagenesis and cloning

5.1.1 RNase E (1-850)

Expression plasmids for wild type RNase E (1-850) and different RNase E (1-529) mutants were available in the lab. Primers for PCR reactions were designed using PrimerX (www.bioinformatics.org/primerx/) (Table 1) and amplification was performed with Pfu Turbo polymerase (NEB). All mutations were incorporated from existing RNase E (1-529) constructs and verified by sequencing.

Two mutations were introduced in the RNase E catalytic domain, D303N/D346N and D346C (section 3.1.1) via a fusion-PCR strategy comprising of two successive PCR reactions (Figure 22). In the first step, two adjacant fragments of RNase E (1-850) were PCR-amplified, i.e. RNase E (1-445) and RNase E (445-850), with complementary overhangs (Figure 22A). For both mutations, the first RNase E fragment (1-445) was amplified from an existing construct of RNase E (1-529) containing these mutations (provided by Dr. Kasia Bandyra, Luisi group). The second RNase E fragment (1-850) was amplified from a wild type construct of RNase E (1-850) (Figure 22). PCR products were resolved on a 1% low melting point agarose gel (Sigma), the bands of interest were excised and, after melting the matrix by incubation at 70°C, mixed in a single second PCR reaction which amplified the whole RNase E (1-850) gene, containing the desired mutations (Figure 22B). The product of the last PCR was digested with SalI and EcoRI (NEB) and resolved on a 1% low melting point agarose gel. The gel band was excised and directly ligated with T4 ligase (NEB) into a pRSFDuet plasmid containing the RhlB gene in one of its multiple cloning sites. The latter had been digested with the same restriction enzymes and dephosphorylated with CIP (NEB) according to the manufacturer's instructions. The primers used for the PCR reactions are depicted in Table 1.

Table 1: **RNase E (1-850) mutagenesis primers.**

Primer name	Primer sequence (5'-3')
RNaseEEcoFor – 1	CCAGGATCCGAATTCGATGAAAAGAATGTTAATCAACG
RNaseE850RevSal – 4	GCACAGGAACAATGGCGTGAACTTCCTGGGTG
NTD_F – 3	CACCCAGGAAGTTCACGCCATTGTTCCTGTGC
NTD_R – 2	CGCAAGCTTGTCGACTTACTCAACAGGTTGC

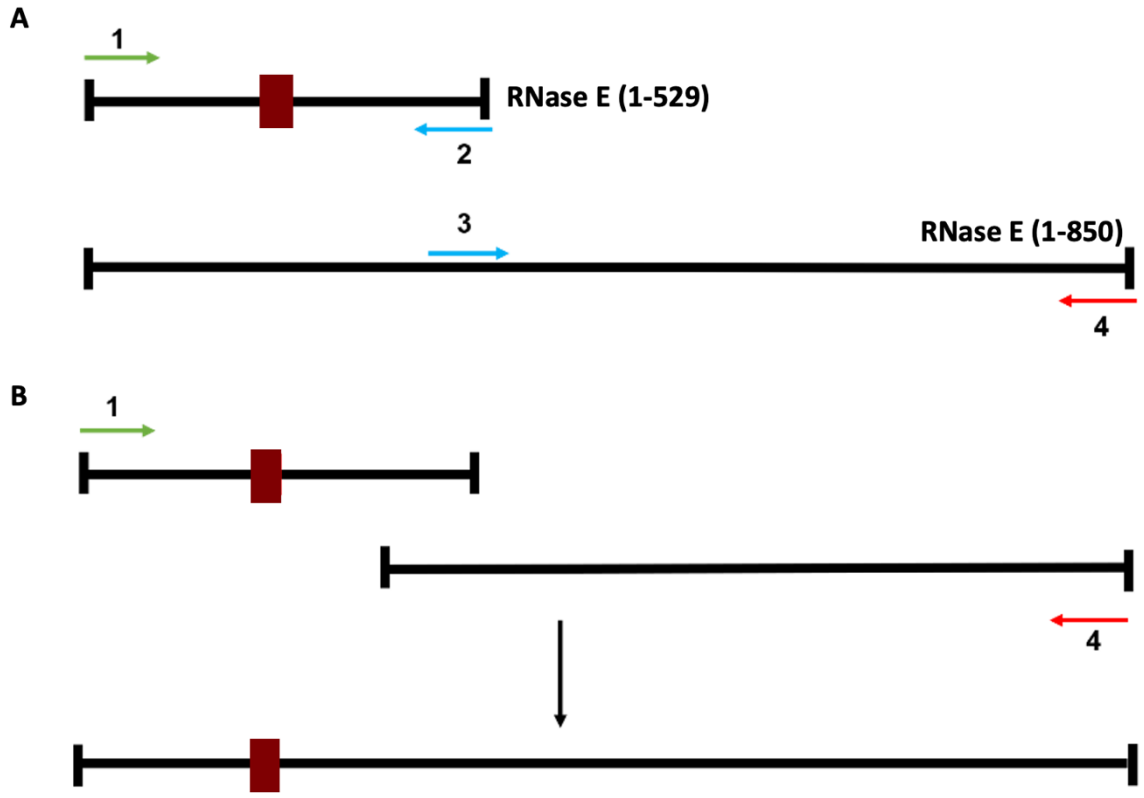


Figure 22: **Schematic of the fusion PCR strategy E.** The numbers indicate primers as depicted in Table 1 and the coloured arrows indicate the direction of the template amplification. The red squares indicate the mutation present in the RNase E (1-529) template. **(A)** Amplification of the two RNase E halves from two different templates (see text). **(B)** Fusion PCR of the two RNase E fragments.

5.1.2 RNase E (1-1061)

One inactivating double mutation, D303N/D346N, was introduced in the catalytic domain of the full-length RNase E (1-1061), following the same protocol as presented in Figure 22. The first fragment was amplified from the same existing RNase E (1-529) construct bearing the D303N/D346N mutations. The second fragment was amplified from a wild type RNase E (1-1061) construct available in the lab. After fusion PCR, the new construct was ligated (T4 ligase) in a pET15 vector which was digested with EcoRI and SalI. The primers used are depicted in Table 2.

Table 2: **RNase E (1-1061) mutagenesis primers**

Primer name	Primer sequence (5'-3')
RnePrsfF – 1	CCAGGATCCGAATTCGATGAAAAGAATGTTAATCAACG
RneIntR – 4	GCACAGGAACAATGGCGTGAACTTCCTGGGTG
RneIntF – 3	CACCCAGGAAGTTCACGCCATTGTTCCTGTGC
RnePrsfR – 2	CGCAAGCTTGTCGACTTACTCAACAGGTTGC

5.2 Protein purifications

5.2.1 *In vivo* assembly of the truncated degradosome

E. coli ENS134-10_eno cells (carrying the pET21b_eno plasmid encoding Enolase) were transformed with pRSF-DUET_rne1-850/rhlb, encoding 6xHis-tagged RNase E (1-850) and RhlB, and used to grow cultures in 2xYT (Formedium) at 37°C supplemented with 15 µg/ml kanamycin and 25 µg/ml carbenicillin. When the OD_{600nm} reached 0.3-0.5, expression was induced with addition of 1mM IPTG and cultures were incubated overnight at 18°C. Cells were harvested by centrifugation with a Beckman JS 4.2 rotor at 5018 g and resuspended in NiA buffer (50 mM Tris-HCl pH 7.5, 1 M NaCl, 100 mM KCl, 10 mM $MgCl_2$, 1 mM TCEP, 0.02% (w/v) ß-DDM), supplemented with a complete EDTA-free protease inhibitor tablet (Roche). 1% (v/v) Triton X-100 was added to help solubilize membrane associated proteins during lysis, along with 1 mM TCEP, 1 mM PMSF, 2 µg/ml DNase I and 0.1% (w/v) ß-DDM. After lysis by cell rupture (Emusliflex-05, Avestin, 5 passes at 10-15 kbar), the lysate was centrifuged at 4 °C (38000 g, 30 min) and loaded on a 5 ml HiTrap chelating HP column (GE healthcare), freshly charged with $NiSO_4$ and equilibrated with NiA buffer. The truncated degradosome was eluted with a linear 0-60% NiB gradient (NiA + 500 mM imidazole). Protein containing fractions were analysed with SDS PAGE, pooled and diluted (1:3) with SP C buffer (50 mM Tris-HCl pH7.5, 0.02% (w/v) B-DDM) before loading on a 1 ml HiTrap SP HP column (GE Healthcare) equilibrated with SP A buffer (50 mM Tris-HCl pH 7.5, 50 mM NaCl, 10 mM KCl, 1 mM TCEP 0.02% (w/v) ß-DDM). The truncated degradosome was eluted with a linear gradient (0-50%) of SP B buffer (50mM Tris-HCl pH 7.5, 2 M NaCl, 10 mM KCl, 1mM TCEP, 0.02% (w/v) ß-DDM). The protein-containing fractions were pooled and concentrated with 100 kDa cutoff Amicon Ultra centrifugation filter (Milipore) according to the manufacturer's protocol. The concentrated sample was then loaded on a Superose 6 10/300 GL size-exclusion column (GE Healthcare) equilibrated with S6 running buffer (50 mM HEPES pH 7.5, 400 mM NaCl, 100 mM KCl, 1 mM TCEP, 0.02% (w/v) B-DDM). Peak fractions were analysed with SDS PAGE and clean fractions were supplemented with 10% v/v glycerol before storage at -80 °C. For subsequent experiments, protein concentrations were determined spectroscopically using a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific) and a λ280nm extinction coefficient of 119990 $M^{-1}cm^{-1}$ for a protomeric truncated degradosome unit.

5.2.2 Endogenous RNA degradosome purification

E. coli NCM3416 cells (kindly provided by the A.J Carpousis lab, LMGM, University of Toulouse, France) were grown in 2XYT medium supplemented with 50 µg/ml kanamycin to an OD_{600nm} of 2.0. Cells were harvested by centrifugation using a Beckman JS 4.2 rotor at 5018 g for 25 min and resuspended in StrepA buffer (50 mM Tris-HCl pH 7.5, 400 mM NaCl, 100 mM KCl, 10 mM $MgCl_2$, 0.01% (w/v) ß-DDM), supplemented with a complete EDTA-free protease inhibitor tablet (Roche). 1% Triton X-100 was added to help solubilize membrane associated proteins during lysis, along with 1 mM TCEP, 2 µg/ml of DNase I and 1 mg/ml Lysozyme (Sigma-Aldrich). After lysis of the cells (Emulsiflex-05, Avestin, 5 passes at 10-15 kbar), the lysate was centrifuged at 4°C (20000 g, 30 min) and loaded on a 1 ml HiTrap HP Strep column (GE healthcare), equilibrated with StrepA buffer. The endogenous RNA degradosome was eluted with a 100% Strep B buffer (50 mM Tris-HCl pH 7.5, 200 mM NaCl, 100 mM KCl, 10 mM $MgCl_2$, 0.02% (w/v) ß-DDM, 2.5 mM Desthiobiotin (Sigma-Aldrich)). The protein-containing fractions were analysed

with SDS-PAGE, pooled and diluted with SP C buffer (1:4), before loading on a 1ml HiTrap Heparin HP column (GE Healthcare) equilibrated with SP A buffer. The RNA degradosome was eluted with a linear gradient (0-50%) of SP B buffer and analysed with SDS-PAGE. Protein containing fractions were pooled, supplemented with 10% (v/v) glycerol and flash frozen with liquid nitrogen for storage at -80°C.

5.2.3 *In vitro* reconstitution of the RNA degradosome

E. coli C43 cells were transformed with pET15_wtrne encoding 6xHis-tagged wild type full-length RNase E, or pET15_NNrne encoding the D304N/D346N inactive full-length RNase E. Cells were grown at 37°C in 2×YT media (Formedium) supplemented with 50 µg/ml kanamycin. At OD_{600nm}=0.6, the cultures were induced with 1 mM IPTG and incubated overnight at 16°C. Untagged Enolase (pET21b_eno), PNPase (pET-Duet_pnp) and RhlB (pRSF_Duet_rhlb) were expressed separately in *E. coli* BL21(DE3) cells according to standard protocol (1mM IPTG induction at OD_{600nm}=0.5, 3-4 hours of expression at 37°C). All cells were mixed together before lysis, corresponding to 6 x 0.5l of RNase E cultures, 1 x 0.5l of PNPase culture, 1 x 0.5l of Enolase culture and 1 x 0.5l of RhlB culture. The cell lysis and protein purification protocols are identical to those described in section 5.2.1. The protein concentration was determined spectroscopically using a NanoDrop ND-1000 spectrophotometer (Thermo Scientific) and a λ280nm extinction coefficient of 218090 $M^{-1}cm^{-1}$ per protomeric RNA degradosome unit.

To reconstitute the full RNA degradosome *in vivo,* pET15_wtrne was transformed to *E. coli* C43 cells to express wild type 6xHis-tagged RNase E. Expression of RNase E and pull-down purification of the RNA degradosome are identical to the protocol described in section 5.2.1.

5.2.4 RNase E catalytic domain

BL21(DE3) cells harbouring the plasmid pRne529-N, which encodes RNase E catalytic domain with N-terminal 6xHis-tag, were grown in 2×YT media (Formedium) supplemented with 100 µg/ml carbenicillin at 37°C. At OD_{600nm}=0.6, the cultures were induced with 1 mM IPTG, cells were harvested by centrifugation after 3 hours of incubation at 37°C. The cells were then resuspended in buffer A (20 mM Tris-HCl pH 7.9, 500 mM NaCl, 5 mM imidazole, protease inhibitor cocktail tablet (Roche)), supplemented with DNase I (1 µg/ml), and passed through an EmulsiFlex-05 cell disruptor (Avestin) at 10-15 kbar for cell lysis (3 passes). The lysate was clarified by centrifugation (4 °C, 30 minutes, 37500 *g*) and loaded on a 5ml HiTrap Chelating HP column (GE Healthcare) charged with $NiSO_4$. Proteins were eluted by an imidazole gradient (buffer A supplemented with 1 M imidazole) and evaluated by SDS–PAGE. Fractions enriched in RNase E catalytic domain were pooled, concentrated to 2 ml with an Amicon Ultra 30,000 MWCO concentrator (Millipore) and loaded on to a Sephadex 200 gel filtration column (GE Healthcare) equilibrated with a running buffer containing 20 mM Tris pH 7.9, 500 mM NaCl, 10 mM $MgSO_4$, 0.5 mM EDTA, 1 mM TCEP, 5% v/v glycerol and a protease inhibitor cocktail tablet (Roche). Eluted fractions were analysed by SDS–PAGE and those containing purified RNase E catalytic domain were flash frozen with liquid nitrogen and stored at -80 °C. The protein concentration was determined spectroscopically using a NanoDrop ND-1000 spectrophotometer (Thermo Scientific) and a λ280nm extinction coefficient of 29005 M-1cm-1 per RNase E (1-529) protomer.

5.2.5 Expression and purification of the recognition core

Expression and purification of the recognition core were carried out following standard protocols described in Bruce *et al.* (2018).

5.3 TCA precipitation and SDS-PAGE

Trichloroacetic acid (TCA) was added to protein samples at 10% (v/v) final concentration from a 50% stock. The samples were then left on ice for 25 min and subsequently spun down at 20 000 g for 45 min (4°C). The supernatant was carefully removed, and the pellet was resolubilised in 20 µl LDS loading dye (95% (v/v) 5xLDS, 5% (v/v) DTT), and analysed with SDS-PAGE. Samples were boiled at 95°C for 3 minutes and loaded on a 12% NuPAGE SDS-gel and run in 1x MOPS Buffer. Gels were run at 150 mV for 60 minutes and stained with Coomassie Blue.

5.4 Sedimentation velocity analytical ultracentrifugation (SV-AUC)

SV-AUC was performed using a Beckman Optima XL-I analytical ultracentrifuge (Biophysics facility, Department of Biochemistry, University of Cambridge). Absorbance data was collected at a wavelength of 280 nm for the truncated degradosome sample, which was concentrated to a final concentration of 4 mg/ml in buffer S6 (50 mM Tris-HCl pH 7.5, 400 mM NaCl, 100 mM KCl, 5 mN $MgCl_2$, 1 mM TCEP, 0.02% ß -DDM, 10% (v/v) glycerol). The samples were spun at 4 °C, 50 000 g for 12 hours. The obtained data were modelled using Sedfit software (Zhao *et al.*, 2013). Buffer viscosity was calculated with SEDNTERP (http://jphilo.mailway.com). AUC experiments were carried out with the help of Dr. Steve Hardwick (BioCem, Department of Biochemistry, University of Cambridge).

5.5 Dynamic light scattering (DLS)

Truncated degradosome samples were diluted in S6 buffer (50 mM Tris-HCl pH 7.5, 400 mM NaCl, 100 mM KCl, 5 mM $MgCl_2$, 1 mM TCEP, 0.02% ß-DDM, 10% (v/v) glycerol) to final concentration of 0.5 mg/ml. A buffer with a lower salt content was also tested for signs of protein aggregation (50 mM Tris-HCl pH 7.5, 100 mM NaCl, 150 mM KCl, 5 mM MgCl2, 1 mM TCEP, 0,02% (v/v) ß-DDM). For the latter, buffer exchange was carried out with a PD MiniTrap G-10 column (GE Healthcare) column. A final sample volume of 70 µl was used for DLS measurements using a *Zetasizer Nano S* (Malvern Panalytical). For each experiment three technical replicates, with 15 runs each, were collected.

5.6 Small Angle X-ray Scattering coupled to Size Exclusion Chromatography (SEC-SAXS)

The truncated degradosome was purified as described in section 5.2.1. The samples were concentrated, flash frozen in liquid nitrogen and sent to beamline B21 at Diamond Lightsource (Harwell campus, Didcot, United Kingdom) for data collection. Prior to freezing 9S rRNA was added to the sample at equimolar ratios (if applicable). The sample was loaded on a Superose 6 Increase 3.2/300 column (GE Healthcare) by a High-Performance Liquid Chromatography (HPLC) instrument (Agilent) directly before elution into the sample detection chamber, where a monochromatic beam illuminated the sample as it flowed through. 2D scattering profiles were radially averaged on the fly at the beamline. For each sample the Radius of Gyration (R_g) was plotted as a function of the elution

frame for optimal selection of a homogeneous sample range in Scatter (biosis.net). Temporal averaging of the selected range and buffer subtraction were carried out in Scatter and in DataSW to generate the final intensity curve (biosis.net, Skumatov & Strelkov, 2015). Further analysis of the intensity curves was performed in Scatter. ALMERGE (Franke *et al.*, 2012) was used to merge intensity curve from different concentrations.

For flexibility assessment of the truncated degradosome, the EOM approach was used (Bernado *et al.*, 2007; Tria *et al.*, 2015). The general procedure for generating a random pool of truncated degradosome models is described in detail in section 3.2.4 of the results. In short, RANCH (Bernado *et al.*, 2007) and FULCHER (Shkumatov e*t al.*, in preparation) were used to generate random pools of truncated degradosome monomer. To generate multiprotein input structures readable by RANCH a python script was written with the help of Dr. Jamie Blaza (Blaza group, YSBL, Department of Chemistry, University of York) (Script 1). The four monomer pools, as generated by RNACH, were then combined randomly with CombinerIM (Shkumatov e*t al.*, in preparation), to generate a pool of approximately 10 million random truncated degradosome structures. The RNase E scaffold domain was modelled as a 'random coil' chain type. After clash score filtering (MOLPROBITY clash score cutoff of 60), a final pool of approximately 5 million truncated degradosome models was used for ensemble modelling. The genetic algorithm *GAJOE* was run 10 times (100 iterations each) to obtain an ensemble of models that best describes the experimental SAXS data.

```
def read_pdb(file_name, new_file):
    raw = open(file_name, 'r')
    pdb = raw.readlines()
    raw.close()
    XX_lst = []
    notXX_lst = []
    for line in pdb:
        if line[13:16] == 'XX ':
            XX_lst.append(line.replace('XX ', 'CA '))
        elif line[13:16] != 'XX ':
            notXX_lst.append(line)
    both = open(new_file,'w')
    for line in notXX_lst:
        both.write(line)
    for line in XX_lst:
        both.write(line)
    both.write('TER')
    both.close()


read_pdb('00001testmon.pdb', 'merged.pdb')
```

Script 1: Script to render RhlB (XX) and Enolase (YY) invisible for RANCH by replacing Ca by XX or YY in the pdb.

5.7 In vitro transcription (IVT)

A 263-bp template encoding *E. coli* 9S rRNA was amplified from pKK233-2 (provided by Dr. Kasia Bandyra, Luisi group) using PCR (primers are depicted in Table 3). The PCR reaction was carried out with Phire Hot Start II polymerase (Finnzymes) according to the manufacturer's protocol. Incorporated in one of the primers was a T7

RNA polymerase recognition sequence (underlined in Table 3). PCR products were checked on a 1% agarose gel and purified with the QIAquick PCR Purification Kit (QIAGEN). The PRC product was used as a template for IVT.

IVT reactions were carried out in 200µl volumes according to standard protocol. Each reaction was set up with 3.5 µg of template, ribonucleotides triphosphate, pyrophosphatase, RNA polymerase T7, spermidine buffer, DTT, RNase OUT and DMSO. The IVT reaction was run at 37 °C for 5 hours, and TURBO DNase I (Invitrogen) was added in the last 30 minutes to digest the DNA template. Transcribed 9S rRNA was then gel-purified on a 4% polyacrylamide gel (LifeScienceProducts). Bands containing RNA were visualised by UV–shadowing and excised. RNA was recovered from gel slices by overnight electroelution at 100 V in 1xTBE buffer using an EluTrap System (Whatman). Lastly RNA was purified with an RNA cleanup kit (Thermo Fisher Scientific) and the concentration was measured with a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific). Purified 9S rRNA was stored in milliQ water at -20°C.

*Table 3: 9S **Template amplification primers for IVT***

Primer	**Primer sequence (5′-3′)**
9SForNew	GTTTTTAATACGACTCACTATAGAAGCTGTTTTGGCGG ATGAGAGAAG
9SRev	CGAAAGGCCCAGTCTTTCGACTGAGC

5.8 GraFix and ultracentrifugation

The Truncated degradosome was purified S6_Hepes buffer (50 mM HEPES pH 7.5, 400 mM NaCl, 100 mM KCl, 5 mM $MgCl_2$, 1 mM TCEP, 0.02% ß-DDM, 10% (v/v) glycerol) and diluted to 5 µM. 200 pmol of truncated degradosome were loaded on a continuous glycerol/glutaraldehyde gradient (10-30% (v/v) glycerol/0-0.15% (v/v) glutaraldehyde) in an ultra-centrifugation tube and spun at 68 000 g for 18h, 4 °C. The continuous gradient was made with a Gradient Master™ base unit (BioComp) according to the manufacturer's protocol, and buffer S6_Hepes was used as a base buffer. The gradient was then carefully fractionated manually into 200 µl fractions (bottom to top) using a peristaltic pump at low flow rate as to not mix fractions in the pump tubing. Each fraction was supplemented with 10mM Tris-HCl pH 7.5 to quench the cross-linking reaction. For each cross-linking reaction, a negative control tube was prepared without glutaraldehyde. The latter were used to detect fractions that contain truncated degradosome via standard dot blots or TCA precipitation. The corresponding fractions from the cross-linking tube were then used for negative stain EM or for cryo-EM (after extensive dialysis to remove glycerol from the buffer).

5.9 Liposome extrusion and truncated degradosome reconstitution

To prepare liposomes that mimic the cytosolic membrane of *E. coli*, 1 mg of DOPE/DOPG was prepared by mixing 70% v/v of DOPE (Avanti Polar Lipids) with 30% v/v of DOPG (Avanti Polar Lipids). Lipids were dissolved in chloroform, mixed in proportion, dried overnight in a desiccator, and purged with Argon. 500 µl of liposome buffer (50 mM Tris-HCl pH 7.5, 50 mM NaCl, 150 mM KCl, 5 mM $MgCl_2$, 1 mM TCEP, 100 mM sucrose) was added to

resuspend lipids, which were left at 4°C for 40 minutes for hydration. Liposomes were prepared by extrusion with a Mini-Extruder (Avanti Polar Lipids, 20 passes) using membranes with 0.05 µm diameter pores (Avanti Polar Lipids). Liposomes were stored in the fridge at 4°C for a maximum of 4 days.

Truncated degradosome reconstitution and ultracentrifugation assays were carried out as described in section 5.8, but without crosslinker. A 10-40% glycerol gradient was set up and ultracentrifugation was carried out at 84 000 g in a SW60-Ti rotor for 12 hours, 4° C. Fractions were precipitated with TCA (section 5.3) and analysed with SDS-PAGE.

5.10 Electrophoretic mobility shift assays (EMSA)

EMSAs were used to assess interactions between the truncated degradosome/recognition core and the 30S ribosomal subunit (provided by Prof. Ben F. Luisi). Binding reactions were performed at room temperature in binding buffer (40 mM Tris-HCl pH 7.5, 100 mM NaCl, 100 mM KCl, 10 mM $MgCl_2$, 5 mM NH_4Cl_2, 1 mM TCEP). Final concentrations of each component are depicted in Figure 13A. After a 15 min of incubation loading buffer (10% (w/v) sucrose) was added to the samples. Samples were loaded on a 0.6% agarose gel (0.6 g of agarose (Sigma-Aldrich) in 100 ml of 24 mM Hepes, 190 mM Glycine, 10 mM $MgCl_2$ at pH 8.3) and run for 6 hours at 4°C in running buffer (24 mM Hepes, 190 mM Glycine, 10 mM $MgCl_2$ at pH 8.3) with an applied electric field of 80 V. Sybr-gold staining (Invitrogen) was used to stain the gel for RNA visualization, Coomassie staining was used for visualisation of protein bands.

5.11 Electron microscopy

5.11.1 Negative stain

Continuous carbon electron microscopy grids (EMResolution) were glow discharged in house (Pelco Easyglow) or by Dr. Lyn Carter (Emitech K100X, PDN, University of Cambridge). 25 µl of protein solution (5 µg/ml-100 µg/ml) was pipetted onto the glow discharged carbon and protein particles were left to adsorb for 30s. GraFix samples were left to absorb for 20 minutes at 4°C. After blotting excess sample with Whatman paper, the grid was washed with 25 µl deionized H_2O for 3s and then blotted dry. This washing procedure was repeated three times, after which 25 µl of 1% (w/v) uranyl acetate (PDN, University of Cambridge) was pipetted onto the grid for staining (30s) followed by blotting. Grids were stored at room temperature. The samples were imaged with a FEI Tecnai F20 electron microscope or a FEI Tecnai Osiris FEG-TEM, both operated at 200kV, at appropriate magnification (spot size 1, Gatan Ultrascan 1000XP detector) (Department of Material Science and Metallurgy, University of Cambridge). Images were collected at 1-5 µm under focus. 2D classification was performed in Relion 2.1 (Scheres *et al.*, 2012).

5.11.2 Cryo-EM, Single particle analysis

Truncated degradosome was purified as described in section 5.2.1. Samples were dialysed against buffer S6 prior to grid preparation. R1.2/1.3 Quantifoil grids (holey carbon) were glow discharged with a Pelco Easyglow glow discharger. 3 µl of sample at 0.1-1 mg/ml was applied to the grids with a Vitrobot Mark IV (FEI), and excess sample

was blotted away prior to freezing in liquid ethane. For GraFix treated samples, glycerol was dialysed out overnight, and the sample was concentrated down to 25 µl prior to grid preparation. A dataset was collected on a 300 kV Titan Krios (FEI) (Nanoscience Centre, University of Cambridge). Data processing was performed in Relion 2.1 (Scheres, 2012).

The RNase E catalytic core (RNase E (1-529)) was concentrated to 10 µM and supplemented with 8 mM CHAPSO right before grid preparation. Grids were prepared following the procedure described above (R2/2 Quantifoil grids). A dataset was collected on a 300 kV Titan Krios (FEI, BioCem, Department of Biochemistry, University of Cambridge), and processed in Relion 3.0 (Zivanov *et al.*, 2018). After 3D refinements, the particles were transferred to cryoSPARC 2.15.0 (Punjani *et al.*, 2017) and subjected to another round of refinement. Prior to 3D variability analysis (3DVA) in cryoSPARC, the particle set was expanded along its pseudo-C2 symmetry axis to allow for symmetry-breaking conformational heterogeneity to be modelled (Punjani & Fleet, 2020). 3DVA was run on a set of 64 000 symmetry expanded particles, solving for 6 modes of variability at a resolution cutoff of 8 Å. Each encoded mode of variability was visualised in 'simple' mode, i.e. the consensus map was morphed along the latent reaction coordinate for each mode. The three top variability modes were retained and are described in section 3.3.2.

The recognition core was co-purified with the 30S small ribosomal subunit via size exclusion chromatography (Superose 6 10/300 GL column, GE Healthcare) (section 3.3.3.1). Fractions were analysed on an SDS-page gel and those containing the 30S-recognition core complex were concentrated using a 15 ml Amicon Ultra 30,000 MWCO concentrator (Millipore). The concentration was measured using a NanoDrop ND-1000 spectrophotometer (Thermo Scientific) and a λ_{260nm} extinction coefficient of 13 394 967 $M^{-1}cm^{-1}$. The complex (0.5 µM) was cross-linked with 0.2% of Glutaraldehyde (Sigma-Aldrich) for 30 minutes at room temperature. The cross-linking reaction was quenched with 10 µl of Tris-HCl pH 7.5. In addition, samples that were not cross-linked were prepared for cryo-EM as well. Glow discharged gold grids (R1.2/1.3 Quantifoil) were prepared on a Vitrobot Mark IV (FEI) and screened with a 200 kV Talos Arctica microscope (BioCem, Department of Biochemistry, University of Cambridge, UK). A data collection was set up for a crosslinked sample on 200 kV Talos Arctica microscope (Falcon III, counting mode). Pre-processing (motion correction, ctf estimation and particle picking/extraction) were carried out in WARP 1.06 (Tegunov & Cramer, 2019). Particle sets were then transferred to cryoSPARC 2.12 (Punjani *et al.*, 2017) for further processing. After several rounds of 2D classification, particles were subjected to careful heterogenous refinement to resolve the density for the recognition core. A negative control dataset, i.e. the 30S small ribosomal subunit on its own, was collected on a Titan Krios (FEI), equipped with a K2 detector (Gatan) as well and processed following the same procedure (BioCem, Department of Biochemistry, University of Cambridge, UK).

5.11.3 Cryo-electron tomography and sub-tomogram averaging

Proteoliposome samples were prepared as described in section 5.9. After dialysis a small aliquot was analysed with SDS-PAGE to check the integrity of the C-terminal domain of RNase E. 0.5-1µl of BSA coated gold fiducials (see section 5.11.3.1) were mixed with 10 µl of proteoliposome sample. The sample was then loaded onto glow

discharged R2/2 Quantifoil grids with a Vitrobot Mark IV (FEI) (BioCem, Department of Biochemistry, University of Cambridge, UK). To allow for thicker vitreous ice, a low blot force was used. Grids were screened with a 200 kV Talos Arctica (BioCem, Department of Biochemistry, University of Cambridge, UK). The 30S-proteoliposome samples and grids were prepared as described above, but the 30S small ribosomal subunit was added to the truncated degradosome-liposome sample in the last two hours of the dialysis step, to allow for interactions with the membrane bound truncated degradosome.

Tilt series were collected on a 200kV Talos Arctica with Tomo4 (FEI) and on a 300kV Titan Krios with Tomo5 (FEI) (BioCem, Department of Biochemistry, University of Cambridge, UK). On the Talos Arctica, images were collected with a Falcon III detector in linear mode (FEI) at a nominal pixel size of 2.29 Å/pix. On the Titan Krios movies were collected (15 frames) with a K3 camera in counting mode (Gatan) at a nominal pixel size of 1.43 Å/pix. All tilt series were collected from -60° to +60° with 3° increment, with image tracking before and after every tilt image, and autofocus for every tilt image. On the Talos Arctica a bidirectional tilt scheme was followed, with a total dose of 99 e/Å^2 per tilt series. On the Titan Krios a single step dose-symmetric tilt scheme was used, with a total dose of 125 e/Å^2 per tilt series.

The tilt series of the membrane-bound truncated degradosome were aligned in Etomo (part of the IMOD package, Kremer *et al.*, 1996). After coarse alignment of the tilt series, a fiducial model was picked and refined manually. Using the refined fiducial model, tilt series were aligned both globally and locally. Eight tomograms (2xbinned) were reconstructed from eight aligned tilt series via weighted back projection (WBP), using a SIRT-like filter for visualization purposes. Alternatively, tomograms were subjected to Nonlinear Anisotropic Diffusion (NAD) filtering in IMOD. To plot the averaged radial density profile of the truncated degradosome on the membrane, spherical liposomes coated with truncated degradosome were centred and cropped in Dynamo (Castaño-Díez *et al.*, 2012, Castaño-Díez *et al.*, 2017). A MATLAB script was written together with Mr. Alister Burt (IBS, Grenoble) to calculate and plot the radial density profile of these sub-tomograms, i.e. the average density profile from the centre of the volume to the edge (Script 2). The script applies an icosahedral symmetry operator to the volume, followed by fine rotational averaging over the X- and Y-axes.

```
%% Crop particle in large enough box for extended degradosome
%% (Use catalogue manager directly on tomograms, pick using projection tool in dtmslice)
%% Extract particles (=cropping) --> big enough to include all density and then some!


%% Load particle
%
v = dread('particle_000001.em');
%% Store useful info (sidelength)
%%
sidelength = size(v,1);
centre_dynamo_convention = (sidelength + 1) / 2;
%% Symmetrise icosahedrally, rotate and sum around X and Y axis to increase sampling
%%in XY-plane
v_ico = dynamo_sym(v,'ico');
v_ico(isnan(v_ico)) = 0;
for l= 1:6
v_ico_rot_{l}=dynamo_shift_rot(v_ico,[0,0,0],[0,-3*l,0]);
v_ico_rot_{l}(isnan(v_ico)) = 0;
v_ico_rot_{l+6}=dynamo_shift_rot(v_ico,[0,0,0],[0,3*l,0]);
v_ico_rot_{l+6}(isnan(v_ico)) = 0;
v_ico_rot_{l+12}=dynamo_shift_rot(v_ico,[0,0,0],[-3*l,0,0]);
v_ico_rot_{l+12}(isnan(v_ico)) = 0;
v_ico_rot_{l+18}=dynamo_shift_rot(v_ico,[0,0,0],[3*l,0,0]);
v_ico_rot_{l+18}(isnan(v_ico)) = 0;
v_ico=v_ico+v_ico_rot_{l}+v_ico_rot_{l+6}+v_ico_rot_{l+12}+v_ico_rot_{l+18};
v_ico(isnan(v_ico)) = 0;
end
v_ico(isnan(v_ico)) = 0;
%% Get images
%%
if floor(centre_dynamo_convention) == centre_dynamo_convention
    idx = centre_dynamo_convention;
else
    idx = [floor(centre_dynamo_convention), floor(centre_dynamo_convention) + 1];
end
im = v_ico(:,:,idx);
if size(im,3) == 2
    im = mean(im,3);
end

im2 = v(:,:,idx);
if size(im2,3) == 2
    im2 = mean(im2,3);
end

%% Get RDF
%%
v_csym = dynamo_sym(v_ico, 100);
v_csym(isnan(v_csym)) = 0;

rdf = v_csym(:,:,idx);
if size(rdf,3) == 2
    rdf = mean(rdf,3);
end

rdf = rdf(:,idx);
if size(rdf,2) == 2
    rdf = mean(rdf,2);
end
```

```
%% different normalisations %%
%
% rdf_raw_923 = -rdf(120:190);
% rdf_mat_norm_923 = -normalize(rdf(120:190));
% %normalisation to membrane max
% rdf_mem_norm_923 = -rdf(120:190)./max(rdf_raw_923);
% plot(rdf_mem_norm_923)

%% original normalisation

rdf = normalize(rdf);

%%
min_rdf = min(rdf);
max_rdf = max(rdf);
range_rdf = max_rdf - min_rdf;
scale_factor = floor(sidelength / range_rdf) * 0.75;
rdf_plot = (-rdf * scale_factor) + mean(idx);

%rdf params%%
rdf_bin2_923 = rdf_plot
%% plot
%
x = linspace(0.5, sidelength+0.5, size(rdf,1)) - 0.5;
linecolor = [1, 0.5, 0];
linewidth = sidelength / 60;
figure;
dview(im); hold on; plot(x, rdf_plot, 'LineWidth', linewidth, 'color', linecolor);
figure;
dview(im2); hold on; plot(x, rdf_plot, 'LineWidth', linewidth, 'color', linecolor);
```

*Script 2: **1D-RDP script.** The 1D-RDP was calculated via icosahedral symmetrization followed by rotational averaging in Dynamo.*

Sub-tomogram averaging of the 30S small ribosomal subunit was carried out following two different strategies. Tilt series collected on the Talos Arctica were aligned in IMOD as described above. Global/local alignments and gold fiducial positions were transferred to emClarity (Himes *et al.*, 2018). EmClarity was used to perform 2D- and 3D-ctf correction, template matching, sub-tomogram extraction and sub-tomogram alignment/averaging. For template matching, a reference map was generated in UCSF Chimera from the 30S structure using pdb 4ADV (Pettersen *et al.*, 2004) and rescaled with emClarity. Template matching and the first rounds of alignment/averaging were carried out with 4x binned sub-tomogram volumes (9.16 Å/pix). The angular searches and shifts of the alignments were refined gradually, and binning was reduced (4.58 Å/pix). Two rounds of TomoCPR were included to refine the tilt-series alignments based on aligned 30S particles. A sub-tomogram average of the 30S small ribosomal subunit was reconstructed at 11.7 Å, which is close to Nyquist (9.16Å), after 17 rounds of alignments. The 30S average was then mapped back into the whole tomogram, preserving its original orientation and replacing the noisy sub-tomogram volumes of the 30S small ribosomal subunit particles. In parallel, TomoSegmemTV was used to model the lipid membranes in the raw tomogram. Together, these two operations present the spatial distribution of the 30S small ribosomal subunits in the tomogram and their position/pose with respect to the lipid membranes, which was visualised in UCSF Chimera.

A different approach was used for the tilt series of 30S-proteoliposomes collected on the Titan Krios. Movies were preprocessed in WARP (Tegunov & Cramer, 2019) (Motion correction and ctf estimation). mdocSpoofer (written by Mr. Alister Burt, https://github.com/alisterburt) was used to generate mdoc metadata files per tilt series. These mdoc files are used by Warp to generate preprocessed and ordered stacks per tilt series. The tilt series were then imported in Dynamo (Castaño-Díez *et al.*, 2012) and aligned with dautoalign4warp (written by Mr. Alister Burt, https://github.com/alisterburt), a MATLAB script that performs automated tilt series alignment in Dynamo. The tilt series alignments were imported into Warp for robust 3D-CTF estimation and tomogram reconstruction. Each tomogram was reconstructed twice, once unfiltered and once deconvolved (for visualization). The same 30S template as described above (scaled to 1.43 Å/pix) was used for template matching in WARP. The template and tomograms were binned to 10.01Å/pix on the fly during template matching, and picked coordinates were reconstructed as sub-tomogram volumes at 5.72Å/pix in Warp, together with a 3D-CTF volume. Metadata and initial poses were recorded in a Relion .star file. The extracted sub-tomogram volumes were then imported into Relion 3.0.8 for 3D classification and 3D refinement. A clean set of particles were then imported in M (Tegunov *et al.*, 2020) for further refinement of the 30S sub-tomogram averages and local refinement of the tilt-series alignments, Ctf refinements, and magnification correction. Sub-tomograms were unbinned, and refined once more with M.

5.11.3.1 Fiducial marker preparation

Gold fiducial markers (10 nm diameter, BBI Solutions, EMC10) were manually coated with bovine serum albumin (BSA) prior to cryo-ET grid preparation. 975 µl of BBI gold beads were buffered with 25 µl of 200 mM NaH_2P0_4 pH 5.0, after which 50 µl of a 5 mg/ml BSA stock (in 5 mM NaH_2P0_4 pH5.0) was added. The mixture was incubated overnight at 4 °C with gentle mixing. BSA-coated gold beads were collected by centrifugation (1h, 50 000 g, 4 °C). The supernatant was removed, and gold beads were resuspended in liposome buffer (50 mM Tris-HCl pH 7.5, 50 mM NaCl, 150 mM KCl, 5 mM $MgCl_2$, 1 mM TCEP). Gold beads were collected again by another round of centrifugation (1h, 50 000 g, 4 °C) and resuspended in ~20 µl of liposome buffer. Gold beads were added to the cryo-ET samples in a 1:10 volumetric ratio (gold beads:sample).

6 References

Adler, M. & Alon, U. (2018) Fold-change detection in biological systems. *Curr. Opin. Syst. Biol.* **8**, 81–89.

Aït-Bara, S., & Carpousis, A. J. (2015). RNA degradosomes in bacteria and chloroplasts: classification, distribution and evolution of RNase E homologs. *Molecular microbiology*, *97*(6), 1021–1135. https://doi.org/10.1111/mmi.13095

Al-Husini, N., Tomares, D. T., Bitar, O., Childers, W. S., & Schrader, J. M. (2018). α-Proteobacterial RNA Degradosomes Assemble Liquid-Liquid Phase-Separated RNP Bodies. *Molecular cell*, *71*(6), 1027–1039.e14. https://doi.org/10.1016/j.molcel.2018.08.003

Al-Husini, N., Tomares, D. T., Pfaffenberger, Z. J., Muthunayake, N. S., Samad, M. A., Zuo, T., Bitar, O., Aretakis, J. R., Bharmal, M. M., Gega, A., Biteen, J. S., Childers, W. S., & Schrader, J. M. (2020). BR-Bodies Provide Selectively Permeable Condensates that Stimulate mRNA Decay and Prevent Release of Decay Intermediates. *Molecular cell*, *78*(4), 670–682.e8. https://doi.org/10.1016/j.molcel.2020.04.001

Banani, S. F., Lee, H. O., Hyman, A. A., & Rosen, M. K. (2017). Biomolecular condensates: organizers of cellular biochemistry. *Nature reviews. Molecular cell biology*, *18*(5), 285–298. https://doi.org/10.1038/nrm.2017.7

Bandyra, K. J., Bouvier, M., Carpousis, A. J., & Luisi, B. F. (2013). The social fabric of the RNA degradosome. *Biochimica et biophysica acta*, *1829*(6-7), 514–522. https://doi.org/10.1016/j.bbagrm.2013.02.011

Bandyra, K. J., Wandzik, J. M., & Luisi, B. F. (2018). Substrate Recognition and Autoinhibition in the Central Ribonuclease RNase E. *Molecular cell*, *72*(2), 275–285.e4. https://doi.org/10.1016/j.molcel.2018.08.039

Bandyra, K. J., & Luisi, B. F. (2018). RNase E and the High-Fidelity Orchestration of RNA Metabolism. *Microbiology spectrum*, *6*(2), 10.1128/microbiolspec.RWR-0008-2017. https://doi.org/10.1128/microbiolspec.RWR-0008-2017

Bayas, C. A., Wang, J., Lee, M. K., Schrader, J. M., Shapiro, L., & Moerner, W. E. (2018). Spatial organization and dynamics of RNase E and ribosomes in *Caulobacter crescentus*. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(16), E3712–E3721. https://doi.org/10.1073/pnas.1721648115

Bernado, P., Mylonas, E., Petoukhov, M.V., Blackledge, M., Svergun, D.I. (2007) Structural Characterization of Flexible Proteins Using Small-Angle X-ray Scattering. *J. Am. Chem. Soc.* **129(17)**, 5656-5664

Blaszczyk, M., Kurcinski, M., Kouza, M., Wieteska, L., Debinski, A., Kolinski, A., & Kmiecik, S. (2016). Modelling of protein-peptide interactions using the CABS-dock web server for binding site search and flexible docking. *Methods (San Diego, Calif.)*, *93*, 72–83. https://doi.org/10.1016/j.ymeth.2015.07.004

Boeynaems, S., Alberti, S., Fawzi, N. L., Mittag, T., Polymenidou, M., Rousseau, F., Schymkowitz, J., Shorter, J., Wolozin, B., Van Den Bosch, L., Tompa, P., & Fuxreiter, M. (2018). Protein Phase Separation: A New Phase in Cell Biology. *Trends in cell biology*, *28*(6), 420–435. https://doi.org/10.1016/j.tcb.2018.02.004

Bresson, S., & Tollervey, D. (2018). Surveillance-ready transcription: nuclear RNA decay as a default fate. *Open biology*, *8*(3), 170270. https://doi.org/10.1098/rsob.170270

Brown, C. J., Takayama, S., Campen, A. M., Vise, P., Marshall, T. W., Oldfield, C. J., Williams, C. J., & Dunker, A. K. (2002). Evolutionary rate heterogeneity in proteins with long disordered regions. *Journal of molecular evolution*, *55*(1), 104–110. https://doi.org/10.1007/s00239-001-2309-6

Bruce, H. A., Du, D., Matak-Vinkovic, D., Bandyra, K. J., Broadhurst, R. W., Martin, E., Sobott, F., Shkumatov, A. V., & Luisi, B. F. (2018). Analysis of the natively unstructured RNA/protein-recognition core in the Escherichia coli RNA degradosome and its interactions with regulatory RNA/Hfq complexes. *Nucleic acids research*, *46*(1), 387–402. https://doi.org/10.1093/nar/gkx1083

Callaghan, A. J., Marcaida, M. J., Stead, J. A., McDowall, K. J., Scott, W. G., & Luisi, B. F. (2005). Structure of Escherichia coli RNase E catalytic domain and implications for RNA turnover. *Nature*, *437*(7062), 1187–1191. https://doi.org/10.1038/nature04084

Carpousis, A. J., Van Houwe, G., Ehretsmann, C., & Krisch, H. M. (1994). Copurification of E. coli RNAase E and PNPase: evidence for a specific association between two enzymes important in RNA processing and degradation. *Cell*, *76*(5), 889–900. https://doi.org/10.1016/0092-8674(94)90363-8

Castaño-Díez, D., Kudryashev, M., Arheit, M., & Stahlberg, H. (2012). Dynamo: a flexible, user-friendly development tool for subtomogram averaging of cryo-EM data in high-performance computing environments. *Journal of structural biology*, *178*(2), 139–151. https://doi.org/10.1016/j.jsb.2011.12.017

Castaño-Díez, D., Kudryashev, M., & Stahlberg, H. (2017). Dynamo Catalogue: Geometrical tools and data management for particle picking in subtomogram averaging of cryo-electron tomograms. *Journal of structural biology*, *197*(2), 135–144. https://doi.org/10.1016/j.jsb.2016.06.005

Chandran, V., Poljak, L., Vanzo, N. F., Leroy, A., Miguel, R. N., Fernandez-Recio, J., Parkinson, J., Burns, C., Carpousis, A. J., & Luisi, B. F. (2007). Recognition and cooperation between the ATP-dependent RNA helicase RhlB

and ribonuclease RNase E. *Journal of molecular biology*, *367*(1), 113–132. https://doi.org/10.1016/j.jmb.2006.12.014

Chao, Y. , Li, L., Girodat, D., Förstner, K.U., Said, N. , Corcoran, C. Śmiga, M., Papenfort, K., Reinhardt, R. Wieden, H.J. *et al.* (2017) In vivo cleavage map illuminates the central role of RNase E in coding and non-coding RNA pathways Mol Cell, *65* , pp. 39-51

Cole, J. L., Lary, J. W., P Moody, T., & Laue, T. M. (2008). Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium. *Methods in cell biology*, *84*, 143–179. https://doi.org/10.1016/S0091-679X(07)84006-4

Commichau, F. M., Rothe, F. M., Herzberg, C., Wagner, E., Hellwig, D., Lehnik-Habrink, M., Hammer, E., Völker, U., & Stülke, J. (2009). Novel activities of glycolytic enzymes in Bacillus subtilis: interactions with essential proteins involved in mRNA processing. *Molecular & cellular proteomics: MCP*, *8*(6), 1350–1360. https://doi.org/10.1074/mcp.M800546-MCP200

Cornejo, E., Abreu, N., & Komeili, A. (2014). Compartmentalization and organelle formation in bacteria. *Current opinion in cell biology*, *26*, 132–138. https://doi.org/10.1016/j.ceb.2013.12.007

Deana, A., & Belasco, J. G. (2005). Lost in translation: the influence of ribosomes on bacterial mRNA decay. *Genes & development*, *19*(21), 2526–2533. https://doi.org/10.1101/gad.1348805

Dendooven, T., Luisi, B.F. (2017) RNA search engines empower the bacterial intranet. Biochem Soc Trans; *45 (4)*: 987–997. doi: https://doi.org/10.1042/BST20160373

Dendooven, T., Luisi, B. F., & Bandyra, K. J. (2020). RNA lifetime control, from stereochemistry to gene expression. *Current opinion in structural biology*, *61*, 59–70. https://doi.org/10.1016/j.sbi.2019.10.002

Dendooven, T., & Lavigne, R. (2019). Dip-a-Dee-Doo-Dah: Bacteriophage-Mediated Rescoring of a Harmoniously Orchestrated RNA Metabolism. *Annual review of virology*, *6*(1), 199–213. https://doi.org/10.1146/annurev-virology-092818-015644

Dendooven, T., Van den Bossche, A., Hendrix, H., Ceyssens, P. J., Voet, M., Bandyra, K. J., De Maeyer, M., Aertsen, A., Noben, J. P., Hardwick, S. W., Luisi, B. F., & Lavigne, R. (2017). Viral interference of the bacterial RNA metabolism machinery. *RNA biology*, *14*(1), 6–10. https://doi.org/10.1080/15476286.2016.1251003

DePristo, M. A., Weinreich, D. M., & Hartl, D. L. (2005). Missense meanderings in sequence space: a biophysical view of protein evolution. *Nature reviews. Genetics*, *6*(9), 678–687. https://doi.org/10.1038/nrg1672

D'Imprima, E., Floris, D., Joppe, M., Sánchez, R., Grininger, M., & Kühlbrandt, W. (2019). Protein denaturation at the air-water interface and how to prevent it. *eLife*, *8*, e42747. https://doi.org/10.7554/eLife.42747

DePristo, M. A., Weinreich, D. M., & Hartl, D. L. (2005). Missense meanderings in sequence space: a biophysical view of protein evolution. *Nature reviews. Genetics*, *6*(9), 678–687. https://doi.org/10.1038/nrg1672

Dressaire C, Picard F, Redon E, Loubi.re P, Queinnec I, Girbal L & Cocaign-Bousquet M (2013) Role of mRNA stability during bacterial adaptation. PLoS One 8: e59059.

Dreyfus M. (2009). Killer and protective ribosomes. *Progress in molecular biology and translational science*, *85*, 423–466. https://doi.org/10.1016/S0079-6603(08)00811-8

Durand, S., Tomasini, A., Braun, F., Condon, C., & Romby, P. (2015). sRNA and mRNA turnover in Gram-positive bacteria. *FEMS microbiology reviews*, *39*(3), 316–330. https://doi.org/10.1093/femsre/fuv007

Felden, B., & Paillard, L. (2017). When eukaryotes and prokaryotes look alike: the case of regulatory RNAs. *FEMS microbiology reviews*, *41*(5), 624–639. https://doi.org/10.1093/femsre/fux038

Franke, D., Kikhney, A.G. & Svergun, D.I (2012). Automated acquisition and analysis of small angle X-ray scattering data. *Nucl Instrum Methods Phys Res A.*;689 52-59. doi:10.1016/j.nima.2012.06.008. CTX:C7149.

Garrey SM, Blech M, Riffell JL, et al. Substrate binding and active site residues in RNases E and G: role of the 5'-sensor. *J Biol Chem*. 2009;284(46):31843-31850. doi:10.1074/jbc.M109.063263

Ghora, B. K., & Apirion, D. (1978). Structural analysis and in vitro processing to p5 rRNA of a 9S RNA molecule isolated from an rne mutant of E. coli. *Cell*, *15*(3), 1055–1066. https://doi.org/10.1016/0092-8674(78)90289-1

Górna, M. W., Carpousis, A. J., & Luisi, B. F. (2012). From conformational chaos to robust regulation: the structure and function of the multi-enzyme RNA degradosome. *Quarterly reviews of biophysics*, *45*(2), 105–145. https://doi.org/10.1017/S003358351100014X

Guzikowski, A. R., Chen, Y. S., & Zid, B. M. (2019). Stress-induced mRNP granules: Form and function of processing bodies and stress granules. *Wiley interdisciplinary reviews. RNA*, *10*(3), e1524. https://doi.org/10.1002/wrna.1524

Hadjeras, L., Poljak, L., Bouvier, M., Morin-Ogier, Q., Canal, I., Cocaign-Bousquet, M., Girbal, L., & Carpousis, A. J. (2019). Detachment of the RNA degradosome from the inner membrane of Escherichia coli results in a global slowdown of mRNA degradation, proteolysis of RNase E and increased turnover of ribosome-free transcripts. *Molecular microbiology*, *111*(6), 1715–1731. https://doi.org/10.1111/mmi.14248

Haimovich, G., Medina, D. A., Causse, S. Z., Garber, M., Millán-Zambrano, G., Barkai, O., Chávez, S., Pérez-Ortín, J. E., Darzacq, X., & Choder, M. (2013). Gene expression is circular: factors for mRNA degradation also foster mRNA synthesis. *Cell*, *153*(5), 1000–1011. https://doi.org/10.1016/j.cell.2013.05.012

Heumann, J. M., Hoenger, A., & Mastronarde, D. N. (2011). Clustering and variance maps for cryo-electron tomography using wedge-masked differences. *Journal of structural biology*, *175*(3), 288–299. https://doi.org/10.1016/j.jsb.2011.05.011

Himes, B. A., & Zhang, P. (2018). emClarity: software for high-resolution cryo-electron tomography and subtomogram averaging. *Nature methods*, *15*(11), 955–961. https://doi.org/10.1038/s41592-018-0167-z

Hui, M. P., Foley, P. L., & Belasco, J. G. (2014). Messenger RNA degradation in bacterial cells. *Annual review of genetics*, *48*, 537–559. https://doi.org/10.1146/annurev-genet-120213-092340

Hunt, A., Rawlins, J. P., Thomaides, H. B., & Errington, J. (2006). Functional analysis of 11 putative essential genes in Bacillus subtilis. *Microbiology (Reading, England)*, *152*(Pt 10), 2895–2907. https://doi.org/10.1099/mic.0.29152-0

Johnson, G. E., Lalanne, J. B., Peters, M. L., & Li, G. W. (2020). Functionally uncoupled transcription-translation in Bacillus subtilis. *Nature*, *585*(7823), 124–128. https://doi.org/10.1038/s41586-020-2638-5

Kastner, B., Fischer, N., Golas, M. M., Sander, B., Dube, P., Boehringer, D., Hartmuth, K., Deckert, J., Hauer, F., Wolf, E., Uchtenhagen, H., Urlaub, H., Herzog, F., Peters, J. M., Poerschke, D., Lührmann, R., & Stark, H. (2008). GraFix: sample preparation for single-particle electron cryomicroscopy. *Nature methods*, *5*(1), 53–55. https://doi.org/10.1038/nmeth1139

Khemici, V., & Carpousis, A. J. (2004). The RNA degradosome and poly(A) polymerase of Escherichia coli are required in vivo for the degradation of small mRNA decay intermediates containing REP-stabilizers. *Molecular microbiology*, *51*(3), 777–790. https://doi.org/10.1046/j.1365-2958.2003.03862.x

Kohler, R., Mooney, R. A., Mills, D. J., Landick, R., & Cramer, P. (2017). Architecture of a transcribing-translating expressome. *Science (New York, N.Y.)*, *356*(6334), 194–197. https://doi.org/10.1126/science.aal3059

Koslover, D. J., Callaghan, A. J., Marcaida, M. J., Garman, E. F., Martick, M., Scott, W. G., & Luisi, B. F. (2008). The crystal structure of the Escherichia coli RNase E apoprotein and a mechanism for RNA degradation. *Structure*, *16*(8), 1238–1244. https://doi.org/10.1016/j.str.2008.04.017

Kremer, J.R., Mastronarde, D.N. and McIntosh, J.N. (1996) Computer visualization of three-dimensional image data using IMOD *J. Struct. Biol.* **116**:71-76

Lehnik-Habrink, M., Newman, J., Rothe, F. M., Solovyova, A. S., Rodrigues, C., Herzberg, C., Commichau, F. M., Lewis, R. J., & Stülke, J. (2011). RNase Y in Bacillus subtilis: a Natively disordered protein that is the functional equivalent of RNase E from Escherichia coli. *Journal of bacteriology*, *193*(19), 5431–5441. https://doi.org/10.1128/JB.05500-11

Leroy, A., Vanzo, N. F., Sousa, S., Dreyfus, M., & Carpousis, A. J. (2002). Function in Escherichia coli of the non-catalytic part of RNase E: role in the degradation of ribosome-free mRNA. *Molecular microbiology*, *45*(5), 1231–1243. https://doi.org/10.1046/j.1365-2958.2002.03104.x

Leroy, M., Piton, J., Gilet, L., Pellegrini, O., Proux, C., Coppée, J. Y., Figaro, S., & Condon, C. (2017). Rae1/YacP, a new endoribonuclease involved in ribosome-dependent mRNA decay in *Bacillus subtilis*. *The EMBO journal*, *36*(9), 1167–1181. https://doi.org/10.15252/embj.201796540

Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H*., et al.* (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta crystallographica. Section D, Structural biology*, *75*(Pt 10), 861–877. https://doi.org/10.1107/S2059798319011471

Lin, Y., Protter, D. S., Rosen, M. K., & Parker, R. (2015). Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Molecular cell*, *60*(2), 208–219. https://doi.org/10.1016/j.molcel.2015.08.018

Iost, I., & Dreyfus, M. (1995). The stability of Escherichia coli lacZ mRNA depends upon the simultaneity of its synthesis and translation. *The EMBO journal*, *14*(13), 3252–3261.

Mackie G. A. (2013). RNase E: at the interface of bacterial RNA processing and decay. *Nature reviews. Microbiology*, *11*(1), 45–57. https://doi.org/10.1038/nrmicro2930

Marcaida, M. J., DePristo, M. A., Chandran, V., Carpousis, A. J., & Luisi, B. F. (2006). The RNA degradosome: life in the fast lane of adaptive molecular evolution. *Trends in biochemical sciences*, *31*(7), 359–365. https://doi.org/10.1016/j.tibs.2006.05.005

Martinez-Sanchez, A., Garcia, I., Asano, S., Lucic, V., & Fernandez, J. J. (2014). Robust membrane detection based on tensor voting for electron tomography. *Journal of structural biology*, *186*(1), 49–61. https://doi.org/10.1016/j.jsb.2014.02.015

Martinez-Sanchez, A., Kochovski, Z., Laugks, U., Meyer Zum Alten Borgloh, J., Chakraborty, S., Pfeffer, S., Baumeister, W., & Lučić, V. (2020). Template-free detection and classification of membrane-bound complexes in cryo-electron tomograms. *Nature methods*, *17*(2), 209–216. https://doi.org/10.1038/s41592-019-0675-5

Mathew, E., Mirza, A., & Menhart, N. (2004). Liquid-chromatography-coupled SAXS for accurate sizing of aggregating proteins. *Journal of synchrotron radiation*, *11*(Pt 4), 314–318. https://doi.org/10.1107/S0909049504014086

Milón, P., & Rodnina, M. V. (2012). Kinetic control of translation initiation in bacteria. *Critical reviews in biochemistry and molecular biology*, *47*(4), 334–348. https://doi.org/10.3109/10409238.2012.678284

Moffitt, J. R., Pandey, S., Boettiger, A. N., Wang, S., & Zhuang, X. (2016). Spatial organization shapes the turnover of a bacterial transcriptome. *eLife*, *5*, e13065. https://doi.org/10.7554/eLife.13065

Morita T, Kawamoto H, Mizota T, Inada T, Aiba H. Enolase in the RNA degradosome plays a crucial role in the rapid decay of glucose transporter mRNA in the response to phosphosugar stress in Escherichia coli [published correction appears in Mol Microbiol. 2007 Jan;63(1):308]. *Mol Microbiol*. 2004;54(4):1063-1075. doi:10.1111/j.1365-2958.2004.04329.x

Murashko, O.N. & Lin-Chao, S. (2017) Escherichia coli responds to environmental changes using enolasic degradosomes and stabilized DicF sRNA to alter cellular morphology. Proc. Natl. Acad. Sci. U. S. A., 114(38):E8025-E8034.

Murray, D. T., Kato, M., Lin, Y., Thurber, K. R., Hung, I., McKnight, S. L., & Tycko, R. (2017). Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. *Cell*, *171*(3), 615–627.e16. https://doi.org/10.1016/j.cell.2017.08.048

Murthy, A. C., Dignon, G. L., Kan, Y., Zerze, G. H., Parekh, S. H., Mittal, J., & Fawzi, N. L. (2019). Molecular interactions underlying liquid-liquid phase separation of the FUS low-complexity domain. *Nature structural & molecular biology*, *26*(7), 637–648. https://doi.org/10.1038/s41594-019-0250-x

Nakane, T., Kimanius, D., Lindahl, E., & Scheres, S. H. (2018). Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. *eLife*, *7*, e36861. https://doi.org/10.7554/eLife.36861

Nicastro, D., Schwartz, C., Pierson, J., Gaudette, R., Porter, M. E., & McIntosh, J. R. (2006). The molecular architecture of axonemes revealed by cryoelectron tomography. *Science (New York, N.Y.)*, *313*(5789), 944–948. https://doi.org/10.1126/science.1128618

Noble, A. J., Dandey, V. P., Wei, H., Brasch, J., Chase, J., Acharya, P., Tan, Y. Z., Zhang, Z., Kim, L. Y., Scapin, G., Rapp, M., Eng, E. T., Rice, W. J., Cheng, A., Negro, C. J., Shapiro, L., Kwong, P. D., Jeruzalmi, D., des Georges, A., Potter, C. S., … Carragher, B. (2018). Routine single particle Cryo-EM sample and grid characterization by tomography. *eLife*, *7*, e34257. https://doi.org/10.7554/eLife.34257

Nott, T. J., Craggs, T. D., & Baldwin, A. J. (2016). Membraneless organelles can melt nucleic acid duplexes and act as biomolecular filters. *Nature chemistry*, *8*(6), 569–575. https://doi.org/10.1038/nchem.2519

Nurmohamed, S., McKay, A. R., Robinson, C. V., & Luisi, B. F. (2010). Molecular recognition between Escherichia coli Enolase and ribonuclease E. *Acta crystallographica. Section D, Biological crystallography*, *66*(Pt 9), 1036–1040. https://doi.org/10.1107/S0907444910030015

O'Reilly, F. J. e*t al.* (2020) In-cell architecture of an actively transcribing-translating expressome. *bioRxiv* 2020.02.28.970111. doi:10.1101/2020.02.28.970111

Pérez-Ortín, J. E., Alepuz, P. M., & Moreno, J. (2007). Genomics and gene transcription kinetics in yeast. *Trends in genetics : TIG*, *23*(5), 250–257. https://doi.org/10.1016/j.tig.2007.03.006

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry*, *25*(13), 1605–1612. https://doi.org/10.1002/jcc.20084

Protter, D., Rao, B. S., Van Treeck, B., Lin, Y., Mizoue, L., Rosen, M. K., & Parker, R. (2018). Intrinsically Disordered Regions Can Contribute Promiscuous Interactions to RNP Granule Assembly. *Cell reports*, *22*(6), 1401–1412. https://doi.org/10.1016/j.celrep.2018.01.036

Punjani, A. & Fleet, D.J. (2020). 3D Variability Analysis: Directly resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM images. *bioRxiv*. doi:10.1101/2020.04.08.032466

Punjani, A., Rubinstein, J. L., Fleet, D. J., & Brubaker, M. A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature methods*, *14*(3), 290–296. https://doi.org/10.1038/nmeth.4169

Py B, Higgins CF, Krisch HM, Carpousis AJ. A DEAD-box RNA helicase in the Escherichia coli RNA degradosome. *Nature*. 1996;381(6578):169-172. doi:10.1038/381169a0

Redko, Y., Aubert, S., Stachowicz, A., Lenormand, P., Namane, A., Darfeuille, F., Thibonnier, M., & De Reuse, H. (2013). A minimal bacterial RNase J-based degradosome is associated with translating ribosomes. *Nucleic acids research*, *41*(1), 288–301. https://doi.org/10.1093/nar/gks945

Richardson J. P. (1991). Preventing the synthesis of unused transcripts by Rho factor. *Cell*, *64*(6), 1047–1049. https://doi.org/10.1016/0092-8674(91)90257-y

Roux, C. M., DeMuth, J. P., & Dunman, P. M. (2011). Characterization of components of the Staphylococcus aureus mRNA degradosome holoenzyme-like complex. *Journal of bacteriology*, *193*(19), 5520–5526. https://doi.org/10.1128/JB.05485-11

Scheres S. H. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. *Journal of structural biology*, *180*(3), 519–530. https://doi.org/10.1016/j.jsb.2012.09.006

Schmid, M., & Jensen, T. H. (2018). Controlling nuclear RNA levels. *Nature reviews. Genetics*, *19*(8), 518–529. https://doi.org/10.1038/s41576-018-0013-2

Schmidli, C., Rima, L., Arnold, S. A., Stohler, T., Syntychaki, A., Bieri, A., Albiez, S., Goldie, K. N., Chami, M., Stahlberg, H., & Braun, T. (2018). Miniaturized Sample Preparation for Transmission Electron Microscopy. *Journal of visualized experiments : JoVE*, (137), 57310. https://doi.org/10.3791/57310

Schmidt, C., Kowalinski, E., Shanmuganathan, V., Defenouillère, Q., Braunger, K., Heuer, A., Pech, M., Namane, A., Berninghausen, O., Fromont-Racine, M., Jacquier, A., Conti, E., Becker, T., & Beckmann, R. (2016). The cryo-EM structure of a ribosome-Ski2-Ski3-Ski8 helicase complex. *Science (New York, N.Y.)*, *354*(6318), 1431–1433. https://doi.org/10.1126/science.aaf7520

Sengoku, T., Nureki, O., Nakamura, A., Kobayashi, S., & Yokoyama, S. (2006). Structural basis for RNA unwinding by the DEAD-box protein Drosophila Vasa. *Cell*, *125*(2), 287–300. https://doi.org/10.1016/j.cell.2006.01.054

Shahbabian, K., Jamalli, A., Zig, L., & Putzer, H. (2009). RNase Y, a novel endoribonuclease, initiates riboswitch turnover in Bacillus subtilis. *The EMBO journal*, *28*(22), 3523–3533. https://doi.org/10.1038/emboj.2009.283

Skou, S., Gillilan, R. E., & Ando, N. (2014). Synchrotron-based small-angle X-ray scattering of proteins in solution. *Nature protocols*, *9*(7), 1727–1739. https://doi.org/10.1038/nprot.2014.116

Shkumatov, A. V., & Strelkov, S. V. (2015). DATASW, a tool for HPLC-SAXS data analysis. *Acta crystallographica. Section D, Biological crystallography*, *71*(Pt 6), 1347–1350. https://doi.org/10.1107/S1399004715007154

Sohlenkamp, C., & Geiger, O. (2016). Bacterial membrane lipids: diversity in structures and pathways. *FEMS microbiology reviews*, *40*(1), 133–159. https://doi.org/10.1093/femsre/fuv008

Stark H. (2010). GraFix: stabilization of fragile macromolecular complexes for single particle cryo-EM. *Methods in enzymology*, *481*, 109–126. https://doi.org/10.1016/S0076-6879(10)81005-5

Strahl, H., Turlan, C., Khalid, S., Bond, P. J., Kebalo, J. M., Peyron, P., Poljak, L., Bouvier, M., Hamoen, L., Luisi, B. F., & Carpousis, A. J. (2015). Membrane recognition and dynamics of the RNA degradosome. *PLoS genetics*, *11*(2), e1004961. https://doi.org/10.1371/journal.pgen.1004961

Sun, M., Schwalb, B., Schulz, D., Pirkl, N., Etzold, S., Larivière, L., Maier, K. C., Seizl, M., Tresch, A., & Cramer, P. (2012). Comparative dynamic transcriptome analysis (cDTA) reveals mutual feedback between mRNA synthesis and degradation. *Genome research*, *22*(7), 1350–1359. https://doi.org/10.1101/gr.130161.111

Svergun, D.I., Barberato, C. & Koch, M.H.J. (1995) CRYSOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates *J. Appl. Cryst.* , **28**, 768-773.

Tejada-Arranz, A., de Crécy-Lagard, V., & de Reuse, H. (2020). Bacterial RNA Degradosomes: Molecular Machines under Tight Control. *Trends in biochemical sciences*, *45*(1), 42–57. https://doi.org/10.1016/j.tibs.2019.10.002

Tegunov, D., & Cramer, P. (2019). Real-time cryo-electron microscopy data preprocessing with Warp. *Nature methods*, *16*(11), 1146–1152. https://doi.org/10.1038/s41592-019-0580-y

Tegunov, D., Xue, L., Dienemann, C., Cramer, P. & Mahamid, J. (2020) Multi-particle cryo-EM refinement with M visualizes ribosome-antibiotic complex at 3.7 Å inside cells. *bioRxiv*. January 2020:2020.06.05.136341. doi:10.1101/2020.06.05.136341

Tria, G., Mertens, H. D. T., Kachala, M. & Svergun, D. I. (2015) Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ* **2**, 207-217

Tsai, Y. C., Du, D., Domínguez-Malfavón, L., Dimastrogiovanni, D., Cross, J., Callaghan, A. J., García-Mena, J., & Luisi, B. F. (2012). Recognition of the 70S ribosome and polysome by the RNA degradosome in Escherichia coli. *Nucleic acids research*, *40*(20), 10417–10431. https://doi.org/10.1093/nar/gks739

Tudek, A., Schmid, M., Makaras, M., Barrass, J. D., Beggs, J. D., & Jensen, T. H. (2018). A Nuclear Export Block Triggers the Decay of Newly Synthesized Polyadenylated RNA. *Cell reports*, *24*(9), 2457–2467.e7. https://doi.org/10.1016/j.celrep.2018.07.103

Voss JE, Luisi BF, Hardwick SW. Molecular recognition of RhlB and RNase D in the Caulobacter crescentus RNA degradosome. *Nucleic Acids Res*. 2014;42(21):13294-13305. doi:10.1093/nar/gku1134

Worrall, J. A., Górna, M., Crump, N. T., Phillips, L. G., Tuck, A. C., Price, A. J., Bavro, V. N., & Luisi, B. F. (2008). Reconstitution and analysis of the multienzyme Escherichia coli RNA degradosome. *Journal of molecular biology*, *382*(4), 870–883. https://doi.org/10.1016/j.jmb.2008.07.059

Zivanov, J., Nakane, T., Forsberg, B., Kimanius, D., Hagen, W.J.H., Lindahl, E. and Scheres, S.H.W. (2018) RELION-3: new tools for automated high-resolution cryo-EM structure determination. *eLife* **7**:e42166.

Zhao, H., Brautigam, C. A., Ghirlando, R., & Schuck, P. (2013). Overview of current methods in sedimentation velocity and sedimentation equilibrium analytical ultracentrifugation. *Current protocols in protein science*, *Chapter 20*, Unit20.12. https://doi.org/10.1002/0471140864.ps2012s71

Zhong, E.D., Bepler, T., Berger, B. & Davis, J.H. (2020) CryoDRGN: Reconstruction of heterogeneous structures from cryo-electron micrographs using neural networks. *bioRxiv*. doi:10.1101/2020.03.27.003871

Appendix I: List of publications

[1] Van den Bossche, A., Hardwick, S., Ceyssens, P-J., Hendrix, H., Dendooven, T., Voet, M., Bandyra, K., Aertsen, A., Noben, J-P, Luisi, B., Lavigne, R. (2016) Structural elucidation of a novel mechanism for the bacteriophage-based inhibition of the RNA degradosome. *eLife*, 5 e16413.

[2] Dendooven, T., Van den Bossche, A., Hendrix, H., Ceyssens, P-J., Bandyra, K., Voet, M., Aertsen, A., Noben, J-P, Hardwick, S., Luisi, B., Lavigne, R. Viral interference of the bacterial RNA metabolism machinery (2017). *RNA Biology*.14(1):6-10. doi:10.1080/15476286.2016.1251003

[3] Dendooven, T., & Luisi, B. F. (2017). RNA search engines empower the bacterial intranet. *Biochemical Society transactions*, *45*(4), 987–997. https://doi.org/10.1042/BST20160373

[4] Pei, X.Y.*, Dendooven, T.*, Sonnleitner, E., Chen, S., Bläsi U, Luisi, B.F. (2019) Architectural principles for Hfq/Crc-mediated regulation of gene expression. *eLife*;8:e43158.

[5] Dendooven, T., Lavigne, R. (2019) Bacteriophage-Mediated Rescoring of a Harmoniously Orchestrated RNA Metabolism. *Annual Review of Virology*; 6:5.1–5.15.

[6] Dendooven, T., Luisi, B.F., Bandyra, K. (2019) RNA lifetime control, from stereochemistry to gene expression. *Current Opinion in Structural Biology*; 61, 59-70.

[7] Oerum, S., Dendooven, T., Catala, M., Gilet, L., Dégut, C., Trinquier, A., Bourguet, M., Barraud, P., Cianferani, S., Luisi, B.F., Ciarán, C. & Tisné, C. (2020) Structures of *B. subtilis* maturation RNases captured on 50S ribosome with pre-rRNAs. *Molecular Cell*; 80 doi.org/10.1016/j.molcel.2020.09.008

[8] Krepl, M., Dendooven, T., Luisi, B.F., Sponer, J. (2020) Structural dynamics at the Hfq-RNA interface can facilitate RNA screening. *J Biol Chem*; In review.

[9] Dendooven, T., Sinha, D. Roesolevá, A., Cameron, T.A. De Lay, N., Luisi, B.F. & Bandyra K.J. (2020) A cooperative PNPase-Hfq RNA carrier complex facilitates bacterial riboregulation. Mol. *Cell;* In review.

[10] Lulla, V.*, Wandel, M.P.*, Bandyra, K.J.*, Dendooven, T.*, Yang, X.*, Doyle, N., Oerum, S., O'Rourke, S., Randow, F., Maier, H.J., Scott, W., Ding, Y., Firth, A.E., Bloznelyte, K. & Luisi, B.F. (2020) Antisense oligonucleotides target a nearly invariant structural element from the SARS-CoV-2 genome and drive RNA degradation.*eLife,,* in review. bioRxiv 10.1101/2020.09.18.304139

[11] Dendooven, T., Paris, G., Shkumatau, A., Kubanska, M., Burt, M., Luisi, B.F. (2020) On the disorganisation of the *Escherichia coli* RNA degradosome. In preparation.

[12] Dendooven, T., Sonnleitner, E., Bläsi U., Luisi, B.F. (2020). Polymorphic ribonucleoprotein folding drives translation-repression in *Pseudomonas aeruginosa*. In preparation.

[13] Wilson, L., Dendooven, T., Hardwick, S., Tryfona, T., Chirgadze, D.Y., Luisi, B. F., Logan, D., Mani, K., Dupree, P. (2020) Cryo-EM structure of bi-domain glycosyltransferase EXTL3 at 2.4 Å resolution reveals the molecular basis for heparan synthase activities. In preparation.

[14] Malecka, E.M., Bassani, F., Dendooven, T., Sonnleitner, E., Rozner, M., Resch, T.A.A., Luisi, B.L., Woodson, S. & Bläsi, U. (2020) Stabilization of Hfq-mediated translational repression by the co-repressor Crc in Pseudomonas aeruginosa. In preparation.

[15] Burt, A., Gaifas, L., Dendooven, T. & Gutsche, I. (2020) Tools enabling flexible approaches to high-resolution sub-tomogram averaging. In preparation.

*these authors contributed equally to this work